PROCEEDINGS OF THE 17TH ANNUAL CONFERENCE ON RESEARCH IN UNDERGRADUATE MATHEMATICS EDUCATION

EDITORS
TIM FUKAWA-CONNOLLY
GULDEN KARAKOK
KAREN KEENE
MICHELLE ZANDIEH

DENVER, COLORADO
FEBRUARY 27 – March 1, 2014

PRESENTED BY
THE SPECIAL INTEREST GROUP OF THE MATHEMATICS ASSOCIATION OF AMERICA (SIGMAA) FOR RESEARCH IN UNDERGRADUATE MATHEMATICS EDUCATION

CITATION: In (Eds.) T. Fukawa-Connelly, G. Karakok, K. Keene, and M. Zandieh, Proceedings of the 17th Annual Conference on Research in Undergraduate Mathematics Education, 2014, Denver, Colorado.

FOREWARD

As part of its on-going activities to foster research in undergraduate mathematics education and the dissemination of such research, the Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education (SIGMAA on RUME) held its sixteenth annual Conference on Research in Undergraduate Mathematics Education in Denver, Colorado from February 27 - March 1, 2014. The conference is a forum for researchers in collegiate mathematics education to share results of research addressing issues pertinent to the learning and teaching of undergraduate mathematics. The conference is organized around the following themes: results of current research, contemporary theoretical perspectives and research paradigms, and innovative methodologies and analytic approaches as they pertain to the study of undergraduate mathematics education. The program included plenary addresses by Dr. Anna Sfard, Dr. Ron Tzur, and Dr. Andrea diSessa and the presentation of over 130 contributed, preliminary, and theoretical research reports and posters. In addition to these activities, faculty, and students contributed to displays on Art and Undergraduate Mathematics Education. The Proceedings of the 17th Annual Conference on Research in Undergraduate Mathematics Education are our record of the presentations given and it is our hope that they will serve both as a resource for future research, as our field continues to expand in its areas of interest, methodological approaches, theoretical frameworks, and analytical paradigms, and as a resource for faculty in mathematics departments, who wish to use research to inform mathematics instruction in the university classroom.

RUME Conference Papers, includes conference papers that underwent a rigorous review by two or more reviewers. These papers represent current work in the field of undergraduate mathematics education and are elaborations of selected RUME Conference Reports. The proceedings begin with the winner of the best paper award and the papers receiving honorable mention. These awards are bestowed upon papers that make a substantial contribution to the field in terms of raising new questions or providing significant or unique insights into existing research programs.

RUME Conference Reports, includes the Poster Abstracts and the Contributed, Preliminary and Theoretical Research Reports that were presented at the conference and that underwent a rigorous review by at least three reviewers prior to the conference. Contributed Research Reports discuss completed research studies on undergraduate mathematics education and address findings from these studies, contemporary theoretical perspectives, and research paradigms. Preliminary Research Reports discuss ongoing and exploratory research studies of undergraduate mathematics education. Theoretical Research Reports describe new theoretical perspectives and frameworks for research on undergraduate mathematics education.

Last but not least, we wish to acknowledge the conference program committee and reviewers, for their substantial contributions to RUME and our institutions, for their support.

Sincerely,

Tim Fukawa-Connelly, RUME Conference Chairperson

Gulden Karakok, RUME Conference Local Organizer

Karen Keene, RUME Program Chair

Michelle Zandieh, RUME Coordinator

PROGRAM COMMITTEE

Program Committee Chair	
Tim Fukawa-Connelly	Drexel University
Local Organizer	
Gulden Karakok	University of Northern Colorado
Karen Allen Keene	North Carolina State University, Raliegh
Jason Belnap	University of Wisconsin Oshkosh
Stacy Brown	California State Polytechnic University, Pomona
Nicole Engelke	West Virginia University
Kyeong Hah Roh	Arizona State University
Elise Lockwood	Oregon State University
Michael Oehrtman	University of Northern Colorado
Vicki Sealey	West Virginia University
Hortensia Soto-Johnson	University of Northern Colorado
Craig Swinyard	University of Portland
Allison Toney	University of North Carolina–Wilmington
Joseph Wagner	Xavier University
Megan Wawro	Virginia Tech
Aaron Weinberg	Ithaca College

CONFERENCE PARTICIPANT LIST

Aditya Adiredja
Valeria Aguirre Holguín
Lara Alcock
Dean Allison
Christine Andrews-Larson
Spencer Bagley
Rabindra Bajracharya
Gregory Baker
Veronica Baker
Brian Beasley
Ahmed Benkhalti
Kristen Bieda
Stuart Boersma
Timothy Boester
Heather Bolles
Andrew Boudreaux
Justin Boyle
Linda Braddy
Aaron Brakoniecki
Stacy Brown
Kelly Bubp
Martha Byrne
Fabiana Cardetti
Debra Carney
Corinne Casolara
Anne Cawley
Danielle Champney
Kritika Chhetri
Warren Christensen
Brian Christopher
Emily Cilli-Turner
Nathan Clements
Darcy Conant
Anna Marie Conner
John Paul Cook
Samuel Cook
RaKissa Cribari
Kathy Czeranko
Jennifer Czocher
Paul Dawkins
Eyob Demeke
Rebecca Dibbs
Andrea diSessa
Marki Dittman
Morgan Dominy
Tevian Dray
Irene Duranczyk
Allison Dorko
Jessica Ellis
Robert Ely
Nicole Engelke Infante
Sarah Enoch
Sam Escolas
Shahram Firouzian
Brian Fisher
Richard Francisco
Tim Fukawa-Connelly
Gillian Galle
David Glassmeyer
Jim Gleason
Erin Glover
Nicholas Gorgievski
Melissa Goss
Anthony Gusman
Lynette Guzman
Samer Habre
Brent Hancock
Kady Hanson
Sarah Hanusch
Nadia Hardy
Andrew Hare
Shandy Hauk
Casey Hawthorne
Charles Hayward
Reinhard Hochmuth
Matthew Inglis
Michael Jacobson
Kendrice James
Melissa James
Wendy James
Anek Janjaroon
Estrella Johnson
Steven Jones
Darja Kalajdzievska
Stephen Kanim
Gulden Karakok
Brian Katz
Karen Keene
Minsu Kim
Jeffrey King
Libby Knott
Mile Krajcevski
Andrew Krause
George Kuster
Yvonne Lai
Sean Larsen
Renee LaRue
Sandra Laursen
Maurice LeBlanc
Linda Leckrone
Misun Lee
Kristen Lew
Elise Lockwood
Alison Lynch
Michael Mack
Ami Mamolo
Corinne Manogue
Lisa Mantini
Frank Marfai
Karen Marrongelle
Jason Martin
Celethia McNeil
Juan Pablo Mejia-Ramos
Kathleen Melhuish
Vilma Mesa
Emily Miller
Melissa Mills
Lee Anna Misterek
Susanna Molitoris Miller
Kyunghee Moon
Kevin Moore
Robert C. Moore
Michelle Morgan
Kristen Lee Murphy
Stacy Musgrave
Emilie Naccarato
Kedar Nepal
Xuan Hien Nguyen
Mike Nicholas
Isaac Nichols
Kristin Noblet
Anderson Norton
Gina Nunez
Michael Oehrtman
Eric Pandiscio
Hyejin Park
Joo-Young Park
Matthew Petersen
Valerie Peterson
Amie Pierone
David Plaxco
Robert Powers
Chris Rasmussen
Douglas Riley
Kitty Roach
Lee Roberson
Jason Samuels
Visala Satyam
Milos Savic
Thorsten Scheiner
Elizabeth Scott-Janda
Vicki Sealey
Alees Seehausen
Annie Selden
John Selden
Natalie Selinski
Anna Sfard
Saliha Sha
Mary Shepherd
Adrian Simpson
John Smith
Mary Smith
Kimberly Sofronas
Hortensia Soto-Johnson
Michael Spannuth
Natasha Speer
Eryn Stehr
Sepideh Stewart
April Strom
Molly Stubblefield
Ashley Suominen
Rebecca Swanson
Craig Swinyard
Jeremy Sylvestre
Jenna Tague
Matthew Thomas
John Thompson
Anna Titova
Marcy Towns
Barbara Trigalet
Jonathan Troup
Jeffrey Truman
Jen Tyne
Ron Tzur
Yusuke Uegatani
Joseph Wagner
Aaron Wangberg
Laura Watkins
Laura Watley
Megan Wawro
Eric Weber
Keith Weber
Aaron Weinberg
Benjamin Wescoatt
Diana White
Nina White
Tim Whittemore
Brianna Wiseman
Paul Yu
Ekaterina Yurasovskaya
Michelle Zandieh
Dov Zazkis
Rina Zazkis

We want to thank Easychair for supporting our conference proposal and proceedings process.

~Tim Fukawa-Connelly

Table of Contents

WHAT IS A PROOF? A LINGUISTIC ANSWER TO AN EDUCATIONAL QUESTION

Keith Weber
Rutgers University

Proof is a central concept in mathematics education, yet mathematics educators have failed to reach a consensus on how proof should be conceptualized. I advocate defining proof as a clustered concept, in the sense of Lakoff (1987). I contend that this offers a better account of mathematicians' practice with respect to proof than previous accounts that attempted to define a proof as an argument possessing an essential property, such as being convincing or deductive. I also argue that it leads to useful consequences for research and teaching.

Key words: Cluster model; linguistics; proof

Defining proof

Proof is an essential part of mathematical practice and a central construct in mathematics education. Contemporary goals of mathematics instruction include having students produce deductive proofs, attend to and critique the proofs of others (NCTM, 2000), distinguish between proofs and invalid arguments (Selden & Selden, 2003), appreciate the importance of proofs and the limitations of other types of evidence (Harel & Sowder, 1998), and understand and learn from the proofs that they observe (Mejia-Ramos et al, 2012). While students are expected to do these things throughout their education (e.g., NCTM, 2000; Schoenfeld, 1994; Stylianides, 2007), proof assumes even greater importance in advanced mathematics course at the university level. In these courses, proofs are a primary way that mathematical content is conveyed to students (Raman, 2004; Weber, 2004) and students' grades in these courses are largely determined by their ability to write proofs about the course content (Weber, 2001).

It is a great irony that there is a consensus amongst mathematics educators that we need to help students understand, appreciate, and construct proofs yet we ourselves cannot agree on what a proof is (Balacheff, 2002; Reid & Knipping, 2010; Weber, 2009). Further, it is undeniable that this has had a negative impact on our research as this has hindered our collective ability to build upon each other's work (Balacheff, 2002; Weber, 2009).

In this theoretical paper, I first observe that previous attempts to define proving in mathematics education have sought to delineate essential properties that all proofs in mathematical practice share and then use these properties to distinguish proofs from non-proofs. I then argue that such properties cannot be found as there is not a consensus amongst mathematicians on what arguments are proofs. I propose instead that proof can be viewed as a *linguistic* or *discursive* category that is not defined analytically; in particular, proof can profitably be viewed as a *cluster concept* in the sense of Lakoff (1987). Finally, I discuss the implications that this characterization of proof can have more mathematics education research and teaching.

Previous attempts to define proof

Defining proof in terms of sense and referent

CadwalladerOlsker (2011) distinguished between two broad ways that proof has been defined. For most of the 20th century, philosophers in the analytic tradition define proof as a formal and syntactic object. More recently, philosophers seeking to provide a descriptive account of mathematical practice have characterized proof as the types of proofs that mathematicians actually read and write. For the most part, mathematics educators have preferred the latter approach.

To frame the distinction between these perspectives, I draw on Frege's themes from the philosophy of language. To understand the meaning of a concept, Frege (1892) distinguished between *Sinn* and *Bedeutung*, or sense and referent. As a simplification, we might regard the sense of a concept is our understanding of what the term means and the referent of a concept is the object(s) signified by the concept. Frege's classic example to distinguish between sense and referent is the claim that "the morning star is the evening star". This was a genuine scientific discovery in the ancient world, yet if we view this sentence in terms of referents, it reduces to the trivial tautology that "Venus is Venus".

Those who define proof as a formal object are providing an unambiguous *sense* of what a proof is. A proof is a linear sequence of well-formed formulae in a formal language with explicitly specified axioms, rules of inference, and conditions for well-formed formulae. Every formula in the proof is either an axiom or derived from previous statements via a rule of inference and this sequence concludes with the theorem being proven. There are several objections to this characterization of proof. The first is that the intersection of the referent for this sense and the proofs that mathematicians actually produce (i.e., what mathematicians practically mean by proof) is small. Consequently, this description of proof cannot really explain how mathematical knowledge is justified or generated since mathematicians rarely produce such proofs (see, for instance, Pelc, 2009). Second, from an educational point of view, this view of proof encourages students and teachers to think of proof as a rule-based technical object, leading students to focus on form over function (e.g., Harel & Sowder, 1998; Schoenfeld, 1988).

The alternative approach is to define proof as the collection of artifacts that mathematicians have labeled as proofs. Here proof is defined by its *referent* rather than its sense and thus has immediate relevance to mathematical practice, and perhaps to classroom practice as well. However, by itself, this description is pedagogically limited. What does it mean to desire that students produce the types of proofs that mathematicians produce? A crude (and useless) interpretation might be to ask students to write their proofs using LaTeX, the text processing system used by most mathematicians. Clearly advice of this nature is not what mathematics educators have in mind. One critique against this recommendation is that proofs being written in LaTex is a *nominal* feature of proof-- i.e., a property that most proofs happen to share-- rather than an *essential* property-- i.e., the property that causes these arguments to be proofs. If mathematics educators are to base descriptions of proof on mathematical practice, as CollandwallerOlsker (2011) urged, then it is imperative that we define proof in terms of essential properties that can inform classroom practice.

Definitions of proof in mathematics education

In this sub-section, I provide a partial list of the ways that different mathematics educators have defined proof. It is common to characterize proof as a *convincing* argument. Most notably, Harel and Sowder (1998) defined a proof as an argument that convinces an individual that an assertion is true[1] and a mathematical proof as an argument that would convince a mathematical community. This is consistent with Davis and Hersh (1981) who defined a proof as an argument that would convince a mathematician who knew the subject, Volminik (1990) who described a proof as an argument that would convince a reasonable skeptic, and Mason, Burton, and Stacy (1982) who called proof an argument that would convince an enemy. Balacheff (1987) also thought of proving as convincing, but he sought to highlight the relative and socially contextual nature of proof by defining it as an argument that would convince a particular community at a particular time.

[1] Harel and Sowder also define a proof as an argument that one would use to persuade others about the truth of an assertion, but this will not be relevant in this paper.

Other researchers have suggested that proving is not merely convincing. Duval (2007) noted that, as opposed to argumentation that seeks to convince an audience of a claim, a proof aims to show that a certain statement is a *logically necessary consequence* of things known to be true. In the same vein, Hoyles and Kuchemann (2002) treated a proof as a *deductive argument*, one that does not admit qualification or rebuttals (in the sense of Toulmin (2003)). Weber and Alcock (2009) characterized proof as an argument within the *representation system of proof*, where there are socially recognized rules (albeit rules that are often ambiguous and implicit) for how assertions can be expressed and how new statements can be derived. Stylianides (2007) made a notable attempt to merge the logical and formal perspectives described in this paragraph with the socially situated nature of proof by saying a classroom argument was a proof if it satisfied three criteria: (a) the argument begins with facts that are both true and known to the classroom community, (b) the argument employs reasoning that is both valid and within the conceptual grasp of the classroom community, and (c) the argument is represented appropriately given the sophistication of the audience. Stylianides (2007) argued that this perspective honored the mathematical integrity of proof while also considering the social context of proof in the classroom.

While these different conceptions of proof are useful in that different researchers are addressing different parts of this multi-faceted construct, this has the undesirable effect of producing inconsistent research findings and limiting the field's ability to develop a cumulative body of knowledge (Balacheff, 2002; Weber, 2009). Consider the surprising observation that undergraduate mathematics majors cannot write proofs (e.g., Moore, 1994; Weber, 2001; Weber & Alcock, 2004) with the claim that young children placed in supportive environment can write proofs (e.g., Maher & Martino, 1996). This can be explained when one realizes the proofs in Maher and Martino's (1996) study are being judged against a different standard than the ones in the undergraduate studies. Stefanie's well-known proof by cases was a convincing deductive argument that convinced her peers and it was a surprisingly sophisticated piece of reasoning from a fifth-grade student. However, Stefanie did not carefully attend to the logical status of the different statements in her proof (which Duval (2007) finds paramount in his conception of proof). She also used everyday language and pictures, rather than the precise verbal-symbolic representation systems that Weber and Alcock (2009) have called for in a proof. I do not think it is wrong to call Stefanie's argument in Maher and Martino's (1996) paper a proof, nor do I think it is problematic for Duval (2007) or Weber and Alcock (2009) to hold the standards for proof that they do. But this analysis does show that researchers investigating proof will have different goals for instruction and different views about student success.

Proving is not convincing

A central claim in this paper is that *none* of the characterizations above offer an accurate description of how proof is practiced in the mathematical community. Every description fails to permit arguments that mathematicians would accept as proofs. For the sake of brevity, I will not provide counterexamples for all proofs, but instead I will illustrate this with *conviction*, as this seems to be the most common way to treat proof in the mathematics community.

Philosophers have used Goldbach's Conjecture to illustrate how mathematicians justifiably gain certainty in statements by empirical evidence. Goldbach's Conjecture asserts that every even number greater than 2 can be written as the sum of two primes. For instance, 4 can be expressed as 2 + 2, 6 can be expressed as 3 + 3, 8 can be expressed as 3 + 5, 10 can be expressed as 3 + 7 or 5 + 5, and so on. This conjecture is unproven and, indeed, mathematicians do not anticipate a proof in the foreseeable future (Baker, 2009). However, there is strong empirical evidence in favor of Goldbach's Conjecture. First, the conjecture has been empirically tested for an extremely large number of cases (4×10^{18}, at the time that this paper was written). Second, we can define a function $G(n)$ to express the number of distinct prime pairs that add up to n; for instance, $G(10) = 2$ because there are exactly two prime pairs that sum to 10, $3 + 7$ and $5 + 5$. $G(n)$ tends to increase as n increases, although not monotonically. For all numbers n greater than 100,000 that have been empirically tested, $G(n)$ is greater than 500. The notion that for some extremely large untested n that $G(n)$ would not only dip below 500 but have a value of 0 (which is what would be needed to disprove the Goldbach Conjecture) seems inconceivable. Most mathematicians believe this empirical evidence is overwhelming and regard Goldbach's Conjecture as true, even in the absence of a proof (e.g., Baker, 2009; Echeverria, 1996; Paseau, 2011). Echeverria claimed, "the certainty of mathematicians about the truth of GC [Goldbach's Conjecture] is complete" (p. 42).

The notion of proving as convincing becomes even more problematic when one recognizes that mathematicians often do not obtain complete conviction from the proofs they read and write. Paseau (2011) made this point nicely as follows:

"That we are in possession of a proof of *p* does not imply we should be certain of *p* […] The proof may be long and hard to follow, so that any flesh-and-blood mathematician should assign a non-zero probability to its being invalid. The longer and more complex the proof, the less secure its conclusions" (p. 143).

Indeed, some argue that empirical evidence can sometimes be *more persuasive* than a deductive proof (e.g., Fallis, 2002; Paseau, 2011); if so, this would seem to completely refute the notion that proving is tantamount to convincing. For further elaboration of this point, see Weber, Inglis, and Mejia-Ramos (2014), where my colleagues and I have an extended discussion of this issue, including a summary of empirical studies that support these points.

None of these perspectives is completely right but all have some merit

While none of the previous definitions accurately discriminates proofs from non-proofs, it is important to note that each provides important insights into the nature of proof. For instance, while it is not the case that a proof is merely a convincing argument, it does seem to be the case that statements that seem obviously true need not be justified in a proof, even if they are difficult to justify deductively (cf., Devlin, 2003). None of the other approaches listed can account for why we have softer justification standards for claims that seem likely to be true than those that do not. Likewise, while some proofs are not purely deductive in the sense that they may admit qualifiers and rebuttals (e.g., computer-assisted proofs may yield false results due to poor programming or computer malfunctions), this approach does explain mathematicians' reluctance to sanction empirical or probabilistic arguments as proofs. Dismissing these perspectives because they do not provide a *perfect* description of proof would be throwing out the proverbial baby with the bathwater.

Why proof cannot be defined in terms of shared properties

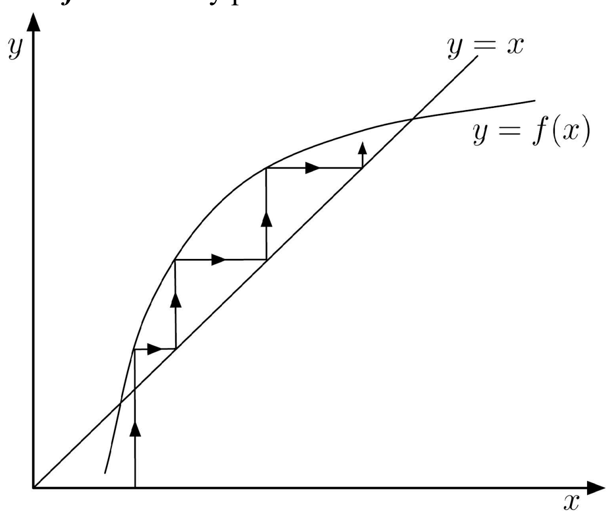
To illustrate the difficulty of finding properties shared by all proofs, consider the following three theorems and proofs that were published in the mathematics literature.

Theorem 1: If n is a number of the form $6k$-1, then n is not perfect.

Proof 1: Assume n is a positive integer of the form $6k$-1. Then $n \equiv -1$(mod 3) and hence n is not a square. Note also that for any divisor d of n, $n = d\left(\frac{n}{d}\right) \equiv -1$(mod 3) implies that $d \equiv -1$(mod 3) and $\left(\frac{n}{d}\right) \equiv 1$(mod 3) or $d \equiv 1$(mod 3) and $\left(\frac{n}{d}\right) \equiv -1$(mod 3). Either way, $d + \left(\frac{n}{d}\right) \equiv 0$(mod 3) and $\sigma(n) = \sum_{d|n, d<\sqrt{n}} d + \frac{n}{d} \equiv 0$(mod 3). Computing $2n=2(6k-1) \equiv 1$(mod 3), we see that n cannot be perfect. (from Holdener, 2002)

Theorem 2: $\pi = \sum_{k=0}^{\infty} \frac{1}{16^k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right)$.

Proof 2: Here is a proof using *Mathematica* to perform the summation.
FullSimplify[TrigtoExp[FullSimplify[
$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right)$]]/.
a_Log[b_]+a_Log[c_]:>a Log[b c]].
π (from Adamchik and Wagon, 1997)
Theorem 3: (Fixed Point Theorem) Let *f(x)* be continuous and increasing on [0, 1] such that $f([0,1]) \subseteq [0,1]$. Let $f_2(x)=f(f(x))$ and $f_n(x)=f(f_{n-1}(x))$. Then under iteration of *f*, every point is either a fixed point or else converges to a fixed point.
Proof 3: The only proof needed is:



(from Littlewood, 1957)

It is difficult to imagine properties shared by these three proofs that distinguish them from non-proofs. In particular, these proofs vary widely in the type of argumentation used, the warrants justifying claims within the proof, their level of transparency, the representation system in which these proofs are couched, and the amount of detail added to justify each claim within these proofs.

A few things should be noted about these proofs. First, some observed that the standards of proof vary across time and mathematical communities (Kleiner, 1991; Rav, 2007). If one accepts this claim, asking whether an argument constitutes a proof is not a well-defined question; one would also need to specify when the argument was given and for what community the argument was intended. Even conceding this point, I do not believe that it is relevant here. These three proofs were intended for a broad mathematical audience-- Proof 1 and Proof 2 appeared in the *American Mathematical Monthly*, an expository journal. Proof 3 appeared in a published volume by J.E. Littlewood (1957). The content of these proofs does not extend beyond the undergraduate curriculum. Specialists' standards of proof do not seem to apply here. There also does not appear to be a reason for why these proofs might be accepted at the time of publication but would be rejected today.

Second, one might argue that these proofs are anomalous. This is true. While Proof 1 was chosen to be fairly representative, I deliberately chose Proof 2 and Proof 3 to be provocative. Still, although these proofs might be atypical, they are nonetheless representative of a class of proofs in mathematics. Proof 2 is analogous to other computer-assisted proofs[2]. Proof 3 is a

[2] The authors noted that, at the time of writing, recent advances had shown that computing the types of sums that they have computed on Mathematica has been and can be certified, as with other computer-assisted computations.

picture proof, which is common in domains such as knot theory and topology, as well as with general mathematical audiences (see Nelsen, 1993).

Proofs and accounting for heterogeneity in mathematical practice*

One might dismiss Proof 2 and Proof 3 as non-proofs. But this is a not a fair argument if we wish to provide a *descriptive* account of proof. The issue at hand is not whether these arguments *ought* to be proofs. The fact is that these *are* proofs as they were published as such in the mathematics literature. A critic could still make the argument that although these proofs were published, this was due to an error on the part of the authors and the editors. Indeed, it should be noted that these proofs *are* regarded as controversial in the mathematics community. In Adamchik and Wagon's (1997) paper in which their proof was presented, the authors admitted that, "Some might even say this is not truly a proof! But in principle, such computations *can* be viewed as proofs" (p. 852). In an experimental study, Inglis and Mejia-Ramos (2009) empirically demonstrated that mathematicians collectively find Proof 3 significantly less convincing than more conventional proofs. Still if these proofs represent mathematicians' errors, it would not be due to *performance error* (e.g., we are not saying that Adamchik and Wagon's (1996) proof contained a typo in the command for *Mathematica*) nor would it be due to a matter of degree (e.g., we would not accept the proof had they used *Maple* instead or checked their results on multiple computers). This would be a stable *epistemological error* and would represent a fundamental difference between how they, and their critics, viewed proofs. I contend these disagreements have an important consequence for how we should understand proof.

Aberdein (2009) coined the term, "proof*", as "species of alleged 'proof' where there is no consensus that the method provides proof, or there is a broad consensus that it doesn't, but a vocal minority or an historical precedent point the other way" (p. 1). As examples of proof*, Aberdein included "picture proofs*, probabilistic proofs*, computer-assisted proofs*, [and] textbook proofs* which are didactically useful but would not satisfy an expert practitioner". As Proof 2 is a computer-assisted proof and Proof 3 is a picture proof, these qualify as proofs*.

Proofs* do not pose a problem for analytic philosophers who attempt to pose *normative judgments* for what should be considered a proof. Recently, there have been arguments that picture proofs, such as Proof 3, are perfectly valid and ought to be on par epistemologically with the more traditional verbal-symbolic proof (for instance, Azzouni (2013); Feferman (2012); and Kulpa (2009)). Granted there may be some mathematicians who disagree, such as the mathematicians in Inglis and Mejia-Ramos' (2009) experimental study, but the proponents of picture proofs can argue that these mathematicians are simply mistaken. That many mathematicians might make epistemologically erroneous judgments is a sociological matter and why individuals do so is a psychological matter. Neither is of normative concern for what constitutes a proof.

However, proofs* *do* pose a problem for philosophers and mathematics educators who wish to describe the proofs that mathematicians actually read and write. Take picture proofs*, for instance. A proposed criteria of proof must either admit some picture proofs* as proofs or claim that all picture proofs* are not. If the former occurred, one could challenge this claim by citing the large number of mathematicians who do not produce such proofs and reject such proofs when they read them. At the very least, such a perspective could not account for the genuine controversy that picture proofs* have caused. If the latter occurred, one could rebut the claim by citing the picture proofs in the published literature as well as the large number of mathematicians (or at least the vocal minority) who accept such proofs. Similar arguments could be made for all types of proofs*. In short, it seems impossible to propose properties that mathematicians believe proofs must share if mathematicians do not themselves agree on what constitutes a proof.

Clustered models of concepts

Lakoff (1987) noted that "according to classical theory, categories are uniform in the following respect: they are defined by a collection of properties that the category members share" (p. 17). This perspective has dominated the way that mathematics educators (and philosophers) have attempted to define proof. However, Lakoff's thesis is that most real-world categories and many scientific categories cannot be characterized this way. In particular, he argued that some categories might be better thought of as *clustered models*, which he defined as occurring when "a number of cognitive models combine to form a complex cluster that is psychologically more basic than the models taken individually" (p. 74). I will argue that mathematical proof should be regarded in the same way

As an illustrative example of a clustered concept, Lakoff considered the category of *mother*. According to Lakoff, there are several types of mothers, including the birth mother, the genetic mother, the nurturance mother (i.e., the adult female caretaker of the child), and the marital mother (i.e., the wife of the father). These concepts are highly correlated-- the birth mother is nearly always the genetic mother and more often than not the caretaker. In the prototypical case, these concepts will converge-- that is, the birth mother will also be the genetic mother, the nurturance mother, and so on. And indeed, when one hears that the woman is the mother of a child, the default assumption is that this woman assumes all of these roles. However, we are aware that this is not always the case.

Lakoff raised two points that will be relevant to this paper. First, there is a natural desire to pick out the "real" definition of mother, or the true *essence of motherhood*. However, Lakoff rejected this essentialist disposition. Different dictionaries list different conceptions of mother as their primary definition. Further, sentences such as, "I was adopted so I don't know who my *real mother* is" and "I am uncaring so I doubt I could be a *real mother* to my child" both are intrinsically meaningful yet define *real mother* in contradictory ways. Second, in cases where there is divergence in the clustered concept of mother (e.g., a genetic but not adoptive mother), compound words exist to qualify the use of mother. Calling one a birth mother typically indicates that she in not the nurturance mother; calling one an adoptive mother or a stepmother indicates that she is not the birth mother.

A clustered model characterization of mathematical proof

The main thesis of this paper is that it would be profitable to consider proof as a clustered concept. The exact models that should form the basis of this cluster should be the matter of debate, but I will propose the following models as a working description to highlight the utility of this approach.

(1) A proof is a *convincing argument* that persuades or ought to persuade a knowledgeable mathematician that a claim is true.

(2) A proof is a *perspicuous argument that provides the reader with an understanding of why a theorem is true.* It provides the reader with an intuitive feeling of necessity.

(3) A proof is a *deductive argument* that shows a theorem is a necessary deductive consequence of previously established claims. A key feature of a proof is that if a proof is correct, there are no potential rebuttals to the argument. The lack of potential rebuttals provides the proof with the psychological perception of being *timeless*. Proven theorems remain proven.

(4) A proof is a *transparent argument where a mathematician can fill in every gap* (given sufficient time and motivation), perhaps to the level of being a formal derivation. In essence, the proof is a blueprint for the mathematician to develop an argument that he or she feels is

complete. Since the mathematician in this description is arbitrary, a proof has the psychological perception of being *impersonal.* Theorems are *objectively* true. In this sense, a proof is a *replicable* demonstration for a given mathematician.

(5) A proof is an *argument within a representation system satisfying communal norms*. That is, there are certain ways of transforming mathematical propositions to deduce statements that are accepted as unproblematic by a community and all other steps need to be justified.

(6) A proof is an *argument that has been sanctioned by the mathematical community*.
Of course, these criteria are not original. Indeed, many are reinterpretations of the definitions of proof given by other mathematics educators and discussed earlier in this paper. What is novel here is claiming that one cannot demarcate proofs from non-proofs by saying that proofs must satisfy some subset of the criteria above.

I argue that each of these more basic models do not, by themselves, characterize proof completely. (1) fails because there are instances in which inductive arguments may be more convincing than deductive arguments (Paseau, 2011). (2) fails because, as Azzouni (2013) noted, lengthy and technical proofs often cannot be meaningfully grasped holistically but are regarded as proofs nonetheless. (3) fails because computer-assisted proofs depend upon the reliability of the software that was used (for a discussion, see Fallis, 2002). (4) fails because, as argued by Fallis (2003), some proofs contain gaps that have not been traversed by any mathematician and that can be extremely difficult to complete.

(5) and (6) are more interesting, since some might argue that these criteria are irrelevant to proofs. Yet if we denied (5) and (6) as being relevant to a proof, we would deny the fact that the standards of proof have clearly changed over time. We would also be compelled to claim that few theorems were proven before 1850 as they would not meet modern day standards of rigor of proof. On the other hand, (5) seems inadequate by itself as a basis for proof because it does not explain why mathematicians permit the transformations that they do. In a sense, (6) appears to be a circular tautology-- the arguments that mathematicians call proofs are the arguments that they sanction as proofs. The point here is that the social act of sanctioning plays a role in the status of an argument beyond logical and psychological considerations. This accounts for the empirical observation that mathematicians are often willing to accept a proof as correct without reading it because it was published in a respectable outlet (see Auslander (2008) and Weber and Mejia-Ramos (2011, 2013), for instance). In short, (5) and (6) describe important social aspects of proof, but do not define proof completely; although proof is necessarily a sociological construct, sociological standards are based in large part on logical and psychological considerations.

If we accept proof to be a clustered concept as defined above, we would expect the following to occur: (a) proofs that satisfied all of these criteria should be uncontroversial, but some proofs that satisfy only a subset of these criteria might be regarded as contentious; (b) compound words exist that qualify proofs that satisfy some of these criteria but not others; (c) it would be desirable for proofs to satisfy all six criteria.

Regarding (a) and (b), Aberdein's (2009) discussion of proofs* supports these points. He explicitly highlighted compound words delimiting the sense that arguments are proofs. Computer-assisted proofs* are not transparent and it is not clear how a mathematician can fill in every gap of the proof. Probabilistic proofs* are not deductive. Picture proofs* are only proofs* in some areas in mathematics-- picture proofs are not controversial in knot theory, among other domains. But in general, picture proofs* are not written in a conventional representation system satisfying the communal norm of being the verbal-syntactic sentences that are accepted as signifying logical propositions. One might add to this list unpublished proofs* that have yet to be sanctioned, heuristic proofs* or incomplete proofs* that require considerable work to fill in gaps, and technical proofs or lengthy proofs that cannot be perceived as a whole (although these proofs seem to be regarded as less controversial than the other types of proof*). Not only do these qualifying compound words exist, but, as Aberdein (2009) argued, there are a significant number of mathematicians who accept such arguments as proofs and a significant number of mathematicians who reject such arguments.

For (c), that it is desirable for proofs to meet all the criteria above, we can consider Dawson's (2006) analysis of why mathematicians re-prove theorems. Proof can be viewed both as a source of knowledge and as justification for knowledge. Some of the reasons for re-proving theorems highlighted by Dawson concern proof as a source for knowledge, such as new proofs can illustrate new methods or extend a previously proven result. But other reasons concern providing a *better justification* for a proven theorem. Dawson notes that one often reproves theorems to remedy perceived gaps and deficiencies in a previous proof. Here, Dawson highlights avoiding non-constructive proofs or proofs that rely on controversial hypotheses, but one could easily extend this argument to avoid proofs that are computer-assisted or are not probabilistic. (It seems intuitively obvious that mathematicians would prefer a proof of the four-color theorem that is not computer-assisted or a deductive proof in lieu of a probabilistic proof*. But the converse would not be true. There would be less demand for a computer-assisted proof* or probabilistic proof* to supplement a conventionally proven theorem). Dawson (2006) further noted that new proofs sometimes "employ reasoning that is simpler, or more perspicuous, than previous proofs" (p. 276), where the new proof is shorter and reduces computations. Dawson also argued that additional proofs provide *confirmation* or *additional conviction* that theorems are true, citing Peirce who claimed that trust in mathematical results stems from the multitude and variety of deductions rather than the conclusiveness of any one demonstration alone. Dawson's analysis aligns with the first four criteria in the proposed proof model above: one reproves theorems to avoid controversial methods, fill in perceived gaps, become more perspicuous, and increase mathematician's conviction. In his paper, Dawson appeared to implicitly view the issue of whether an argument was a proof or whether a theorem was proven as a binary question. However, if we accept a cluster model of proof, we see that some arguments are better representative of the concept of proof than others and the issue of whether a theorem is truly proven might be regarded as a matter of degree. If so, we might say that sometimes theorems are not so much re-proven as *more proven*.

Significance

What this means for mathematics education research

There is a natural desire amongst mathematics educators to specify unambiguous criteria to distinguish proofs from non-proofs. However, if we accept that proof is a cluster concept, this desire cannot be obtained (any more than we can develop a set of rules that would tell us if a study in mathematics education was sufficiently rigorous). I describe what this might mean, first in terms of theory and then in experimental design.

With respect to theory, current definitions of proof generally privilege one aspect of the cluster concept while minimizing or ignoring the other aspects of the cluster concept. This is problematic for the following reason. While the components of the proof cluster are correlated with each other (e.g., an argument that is more explanatory or more based on deductive reasoning will usually be more convincing or more transparent), this is not the case when we take any individual criteria to the extreme. If proving is *only* about convincing, then demonstrations using Geometer's Sketchpad should constitute proofs, as these demonstrations are completely convincing both to students and mathematicians (de Villiers, 2004). Indeed, students likely will find Sketchpad demonstrations to be *more* convincing to than deductive proofs because proof is still a foreign notion to this population of students. Similarly, if proof is *only* about explanation, then there does not seem to ever be a reason to move beyond picture proofs. (Even if mathematics educators believe that picture proofs *should* be proofs, we would still want students to be aware that, and understand why, other mathematicians disagree). Rather than create contrived explanations for why certain types of proofs* are not convincing or explanatory (e.g., we can't trust Sketchpad since the computer might malfunction), we can instead say that the best proofs satisfy multiple aims and the proofs* in question only meet a subset of these aims.

In terms of experimental design, a cluster conception of proof implies researchers must be cautious when participants (students, mathematicians) are required to make dichotomous judgments on whether a specific argument constitutes a proof. First, while researchers frequently act on the assumption that there is a right answer, this might not be the case, especially if a proof* is used. This renders normative judgments about participant's responses to proof evaluation tasks problematic. Second, it may be the case that the argument is a proof in some respects but not others; compelling the participant to make a binary choice could provide a misleading view of what the participant actually believed.

I illustrate this dilemma by re-interpreting a finding from studies that my colleagues and I conducted (e.g., Inglis & Alcock, 2012; Weber, 2008; for a synthesis of these studies, see Inglis et al., 2013). In these studies, the researchers sought to investigate how mathematicians determined if a proof was correct by using materials from Selden and Selden's (2003) classic study on proof validation. One argument, which Selden and Selden labeled "the real deal", aimed to show that "if 3 divides n^2, then 3 divides n". The proof began with the lines, "suppose to the contrary that n is not a multiple of 3. We will let $3k$ be a positive integer that is a multiple of 3 so that $3k+1$ and $3k+2$ are not multiples of 3". Despite the somewhat awkward presentation of the beginning of this proof by contradiction, the logic in the proof was essentially correct. It turned out that the mathematicians in our study disagreed on the validity of this proof; 14 mathematicians judged the proof to be valid while six judged the proof to be invalid. We concluded that mathematicians may have different standards when evaluating proofs (Inglis et al, 2013).

If we view proof as a cluster concept, then one might say that there may have been no disagreement between the mathematicians at all. Rather, what we observed was the consequence of asking an invalid decontextualized question. The argument was a proof in the sense that it was convincing and deductive, but not a proof in the sense of it being couched in an appropriate representation system. It might not be that mathematicians have different standards of proof-hood so much as they had a different interpretation of an artificial decontextualized question when they were forced to give a binary response (in the same way that we might give different responses if asked if a birth mother is the real mother of an adopted child). Selden and Selden (2003) found it problematic that three of the eight mathematics majors participating in this study initially could not make a judgment on whether this argument was a proof, but indeed, given the mathematicians' repsonses, that may have been the most appropriate judgment.

The main point here is that it might not be best to present proof validation tasks as a binary choice in research settings. If we asked participants to say in what respects they thought the argument was a proof and in what respects it was not, we would gain both a more accurate and more detailed understanding of how they thought about proof.

What this might mean for mathematics instruction

Clearly a theoretical analysis cannot offer direct consequences for instruction. However, the ideas in this paper can offer a starting point for future teaching experiments. If proof is a mathematical concept or an analytic concept, then it is important for students to be able to know, interpret, and apply the definition of proof. But if proof is a linguistic or discursive concept, then students might best learn the meaning of proof the way they learn any new word-- based on enculturation. A key role of the instructor is call students' attentions to critical aspects of arguments, both aspects that are proof-like and non-proof-like (and not just formal criteria). Also, in classroom discussions, once an argument is sanctioned as a proof, the debate is often settled and the claim is sanctioned as true. What this paper suggests is that even if an argument passes a threshold and is judged as a proof, there is often still room for improvement. Gaps can be filled in, dubious methods can be replaced by more sound ones, arguments can be made more perspicuous and comprehensible, and greater agreement among classmates can be reached. Proofs can be made better and, in doing so, learning opportunities can be created. Whether these suggestions will actually improve students' appreciation and understanding of proof should be treated as an open question and investigated in a classroom study. Still, the cluster does offer some interesting and unusual ways to treat proof in undergraduate mathematics classrooms.

References

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science*, 14, 1-8.

Adamchik, V. & Wagon, S. (1997). A simple formula for π. *American Mathematical Monthly*, 104, 852-855.

Auslander, J. (2008). On the roles of proof in mathematics. In B. Gold & R. A. Simons (Eds.) *Proofs and other dilemmas: Mathematics and philosophy* (pp. 61-77). Washington, DC: Mathematical Association of America.

Azzouni, J. (2013). The relationship of derivations in artificial languages to ordinary rigorous mathematical proof. *Philosophia Mathematica*, 21, 247-254.

Balacheff, N. (1987). Processus de preuves et situations de validation. *Educational Studies in Mathematics*, 18(2), 147-176.

Balacheff, N. (2002). The researcher epistemology: A deadlock from educational research on proof. In F.L. Lin (ed.) *2002 International Conference on Mathematics Education-Understanding proving and proving to understand*. NCS and NSUT: Taipei.

CadwalladerOlsker, T. (2011). What do we mean by mathematical proof? *Journal of Humanistic Mathematics*, 1, 33-60.

Dawson, J. (2006). Why do mathematicians re-prove theorems? *Philosophia Mathematica*, 14, 269-286.

Davis, P. and Hersh, R. (1981). *The mathematical experience*. New York: Viking Penguin Inc.

de Villiers, M. (2004). The role and function of quasi-empirical methods in mathematics. *Canadian Journal of Science, Mathematics, and Technology Education*, 4, 397-418.

Devlin, K. (2003). The shame of it. From *Devlin's Angle*, an on-line publication of the Mathematical Association of America. Downloaded from: http://www.maa.org/devlin/devlin_archives.html. Last downloaded: March 7, 2014.

Duval, R. (2007). Cognitive functioning and understanding the mathematical processes of proof. In P. Boero (ed.) *Theorems in schools: From history, epistemology, and cognition to classroom practice*. Rotterdam: Sense Publishers.

Echeverria, J. (1996). Empirical methods in mathematics. A case study: Goldbach's Conjecture. In G. Munevar (Ed.) *Spanish Studies in the Philosophy of Science*. (pp, 19-56). Dordrecht: Kluwer.

Fallis, D. (2002). What do mathematicians want? Probabilistic proofs and the epistemic goals of mathematicians. *Logique et Analysese*, 45, 373-388.

Fallis, D. (2003). Intentional gaps in mathematical proofs. *Synthese*, 134, 45-69.

Feferman, S. (2012). And so on…: reasoning with infinite diagrams. *Synthese*, 186, 371-386.
Frege, G. (1892). On *Sinn* and *Beteutung*. In the Frege reader (Ed. Michael Beaney (London: Blackwell, 1997) pp. 151-171.
Harel, G. & Sowder, L. (1998). Students proof schemes. *Research in Collegiate Mathematics Education, 3*, 234-282.
Holdener, J. (2002). A theorem of Touchard on the form of odd perfect numbers. *American Mathematical Monthly*, 109, 661-663.
Hoyles, C. & Kuchemann, D. (2002). Students' understanding of mathematical implication. *Educational Studies in Mathematics*, 51, 193-223.
Inglis, M. & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43(4), 358-390.
Inglis, M. & Mejia-Ramos, P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27, 25-50.
Inglis, M., Mejia-Ramos, J.P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science*, 5, 270-282.
Kleiner, I. (1991). Rigor and proof in mathematics. A historical perspective. *Mathematics Magazine*, 64, 291-314.
Kupla, Z. (2009). Main problems with diagrammatic reasoning. Part 1: The generalization problem. *Foundations of Science*, 14, 75-96.
Lakoff, G. (1987). *Women, fire, and dangerous things.* Cambridge University Press: Cambridge.
Littlewood, J. (1953). *A mathematician's miscellany*. London: Methuen.
Maher, C. A., & Martino, A. M. (1996). The development of the idea of mathematical proof: A 5-year case study. *Journal for Research in Mathematics Education,* 194-214.
Mason, J., Burton, L., & Stacey, K. (1982). *Thinking mathematically*. London: Addison Wesley.
Mejia-Ramos, J.P., Fuller, E., Weber, K., Samkoff, A. & Rhoads, K. (2012). A model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79, 3-18.
National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.
Nelsen, R. (1993). *Proofs without words.* Mathematical Association of America: Washington, D.C.
Paseau, A. (2011). Mathematical instrumentalism, Godel's theorem, and inductive evidence. *Studies in History and Philosophy of Science*, 42, 140-149.
Pelc, A. (2009). Why do we believe theorems? *Philosophia Mathematica*, 17, 84-94.
Raman, M. (2004). Epistemological messages conveyed by three high school and college mathematics textbooks. *Journal of Mathematical Behavior*, 23, 389-404.
Rav, Y. (2007). A critique of a formal-mechanist version of the justification of arguments in mathematicians' proof practices. *Philosophia Mathematica*, 15, 291-320.
Reid, D. & Knipping, C. (2010). *Proof in mathematics education*. Sense Publishers: Rotterdam.
Schoenfeld, A.H. (1988). When good teaching leads to bad results: The disasters of 'well-taught' mathematics courses. *Educational Psychologist*, 23, 145-166.
Schoenfeld, A.H. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, 13(1), 55-80.
Selden, A. & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 36(1), 4-36.

Stylianides, A. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 38, 289-321.
Toulmin, S. (2003). *The uses of argument*. Cambridge University Press: Cambridge.
Volminik, J. (1990). The nature and role of proof in mathematics education. *Pythagorus*, 23, 7-10.
Weber, K. (2001). Student difficulties in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, 48(1), 101-119.
Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, 23(2), 115-133.
Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, 39, 431-459.
Weber, K. (2009). Review of Boero's Theorems in schools. *Mathematical Thinking and Learning*, 11, 289-294.
Weber, K. & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, 56, 209-234.
Weber, K. & Alcock, L. J. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D.A. Stylianou, M.L. Blanton & E. Knuth (Eds.), *Teaching and Learning Proof Across the Grades: A K-16 Perspective* (pp. 323-338). New York, NY: Routledge.
Weber, K., Inglis, M., & Mejía-Ramos, J.P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist*, 49, 36-58.
Weber, K. & Mejía-Ramos, J.P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329-344.
Weber, K. & Mejia-Ramos, J.P. (2013). On influence of sources in the reading of mathematical text: A reply to Shanahan, Shanahan, and Misischia. *Journal of Literacy Research*, 45, 87-96.

GENERALIZING CALCULUS IDEAS FROM TWO DIMENSIONS TO THREE: HOW MULTIVARIABLE CALCULUS STUDENTS THINK ABOUT DOMAIN AND RANGE

Allison Dorko & Eric Weber
Oregon State University

We analyzed multivariable calculus students' meanings for domain and range and their generalization of that meaning as they reasoned about domain and range of multivariable functions. We found that students' thinking about domain and range fell into three broad categories: input/output, independent/dependent variables, and/or as attached to specific variables. We used Ellis' (2007) actor-oriented generalizations framework to characterize how students generalized their meanings for domain and range from single-variable to multivariable functions. This framework focuses on the process of generalization – what students see as similar between ideas in multiple contexts. We found that students generalized their meanings for domain and range by relating objects, extending their meanings, using general principles and rules, and using/modifying previous ideas. Our results about how students understand and generalize the concepts of domain and range imply that the domain and range of multivariable functions is a topic instructors should explicitly address.

Key words: Calculus, function, generalization

Introduction

This paper focuses on (a) how multivariable calculus students think about domain and range in two and three dimensions and (b) how they generalize their meaning of domain and range from single to multivariable functions. We have two foci because how students generalize their ideas cannot be studied without first identifying what those ideas are. While it is clear to experts that multivariable calculus topics are natural extensions of single-variable calculus topics, how students come to see the relationship between ideas like function and rate of change in single and multivariable contexts is not well understood. Though some recent advances have been made with regard to student thinking about these ideas, these studies are only preliminary (Kabael, 2011; Martinez-Planell & Trigueros, 2013; Trigueros & Martinez-Planell, 2010; Yerushalmy, 1997). Additionally, while there is a large body of knowledge about how students understand various single-variable calculus concepts, far fewer studies exist regarding students' understanding of topics in multivariable calculus. For instance, there is a wide body of knowledge about students' understanding of derivatives of single-variable functions (Asiala, 1997; Orton, 1983; Zandieh, 2000), but not much about students' understanding of derivatives of multivariable functions. This scarcity creates two issues: one, we do not know how students in multivariable calculus think about the concepts presented to them and two, we do not understand how they develop those understandings through the process of generalization.

Gaining insight into these two issues is crucial to many STEM fields, as most mathematics used in the real world involves functions of many variables. For instance, in thermodynamics, energy is a function of pairs of pressure, temperature, volume, and entropy; in engineering, density may be a function of x, y, and z. If the mathematics STEM students are to use involves functions of many variables, it makes sense to study how students understand these functions so that instructors can use that knowledge to address specific difficulties and misconceptions. It is likely that students' understanding of single-variable functions plays a role in their understanding of multivariable functions. Thus this study aimed at not only describing one particular aspect of students' multivariable function understanding, but how that thinking relates to their prior knowledge: in short, what they *see as similar* between domain and range of single and multivariable functions. More broadly, knowing how students generalize in mathematics is useful for instruction in that we can better build on students' prior knowledge and exploit the connections they naturally see between mathematical ideas.

We use domain and range as a 'case study' of how students generalize the meaning of a concept learned with single-variable functions to its meaning for multivariable functions. While domain and range appear in initial instruction about functions, they receive little to no attention in multivariable calculus. For instance, McCallum et al. (2009) do not discuss the domain and range of a function at all. Rowgawski (2008) and Thomas (2010) define and give a few examples of the domains and ranges of multivariable functions. None of these standard texts, however, talk about domain and range in terms of inputs and outputs or independent and dependent quantities, as is commonly done in algebra. Thus most of our subjects had not thought about domain and range in three dimensions, and we were able to observe their initial fits and starts with the ideas and observe detailed and sudden generalizations. This paper centers on the following three organizing themes:

1. What meanings do multivariable calculus students have for domain and range in two dimensions?
2. What meanings do multivariable calculus students have for domain and range in three dimensions?
3. How do multivariable calculus students generalize the concept of domain and range from two dimensions to three dimensions?

Background Literature

There are few articles that discuss students' understanding of domain and range. We searched for articles about students' understanding of domain and range, and when that yielded nothing, we switched to associated terms like 'function machines,' 'input and output,' and 'students' notion of variable'. We searched for 'independence and dependence' in both function literature and statistics education literature. None of these searches resulted in articles that explicitly discuss domain and range, though there are some findings in the function literature related to students' understanding of functions that are relevant to the present study. For instance, one way to define domain and range is the set of inputs and outputs of the function, respectively. According to Oehrtman, Carlson, and Thompson (2008), thinking about a function in terms of an input and corresponding output is the beginning of a robust function conception. Monk (1994) found that most calculus students have developed this pointwise view of function but fewer develop an across-time view of function, in which students' conception of function progress to thinking about the function for infinitely many values and understanding how the a change in one variable affects the other(s). That is, a robust function conception involves not only the ability to pair an input with an output, but an understanding of the relationship between quantities. Confrey and Smith (1995) say the beginning of this understanding occurs as students form connections between values in a function's domain and range. However, as function is introduced in algebra and/or precalculus, the functions instructors ask students to reason about are single-variable functions. How students build an understanding of multivariable functions is not known. Our investigation of students' meanings for domain and rage contributes to the function literature by documenting how students think about domain and range of single- and multivariable functions, and how they generalize the ideas of domain and range.

Generalization

We chose to study this sense making in terms of generalization because the ideas in multivariable calculus are connected to those in single-variable calculus (and, in the case of domain and range, to ideas from algebra), and it is widely believed that students use their prior knowledge in making sense of new topics. More specifically, the ideas in multivariable and single variable calculus are similar and students are likely to pick up on similarities such as terms (e.g., function, domain, range, variable) and symbols (e.g. notations like $f(x)$ and $f(x,y)$; integral symbols). Studying the extension from single to multivariable calculus allows us to see the nature of the connections students make and how they make them. Though there have been many studies about generalization in algebra (e.g. Amit & Klass-Tsirulnikov, 2005; Carpenter & Franke, 2001; Cooper & Warren, 2008; Ellis, 2007), these studies are largely about generalizing patterns, and there are fewer studies of generalization of undergraduate mathematics topics, or studies of the generalization of meaning. As generalization is a critical component of mathematical thinking (Amit & Klass-Tsirulnikov, 2005; Lannin, 2005; Mason, 1996; Peirce, 1902; Sriraman, 2003; Vygotsky, 1986), it is important to extend knowledge of how students generalize in higher mathematics, and in particular how they generalize conceptual meanings.

Theoretical Framework

We studied generalization from an actor-oriented perspective. The actor-oriented perspective attends to what students see as similar in mathematical situations. This is in contrast to an observer-oriented perspective in which students' ideas are compared to what an expert would see as similar across situations. Such perspectives often find that students cannot or do not generalize ideas from one setting to another, and focus on the product – the final general rule or principle – as opposed to the generalization process itself. The actor-oriented perspective allows us to privilege students' perceptions of similarity, and thus their generalization process, even if their perceptions are not necessarily consonant with what an expert would see as similar. We follow Ellis (2007) and Lobato (2003) in thinking about generalization as "the influence of a learner's prior activities on his or her activity in novel situations" (Ellis, 2007, p. 225). This was a useful lens for looking at how students viewed domain and range, a topic they had experienced prior with single-variable functions, in the novel situation of multivariable functions.

Our corresponding analytic framework is Ellis' (2007) generalizations taxonomy. The taxonomy distinguishes between *generalizing actions*, or "learners' mental acts as inferred through the person's activity and talk" (Ellis, 2007, p. 233) and *reflection generalizations*, which are students' public statements about a property or pattern common to two situations. Generalizing actions include relating, searching, and extending (Figure 1). Reflection generalizations include identifications and statements, definitions, and influence (Figure 2). We used this framework to analyze how students generalized their meanings for domain and range.

GENERALIZING ACTIONS		
Type I: Relating	1. Relating situations: The formation of an association between two or more problems or situations.	*Connecting Back*: The formation of a connection between a current situation and a previously-encountered situation.
		Creating New: The invention of a new situation viewed as similar to an existing situation.
	2. Relating objects: The formation of an association between two or more present objects.	*Property*: The association of objects by focusing on a property similar to both.
		Form: The association of objects by focusing on their similar form.

<table>
<tr><td rowspan="4">Type II: Searching</td><td>1. Searching for the Same Relationship: The performance of a repeated action in order to detect a stable relationship between two or more objects.</td></tr>
<tr><td>2. Searching for the Same Procedure: The repeated performance of a procedure in order to test whether it remains valid for all cases.</td></tr>
<tr><td>3. Searching for the Same Pattern: The repeated action to check whether a detected pattern remains stable across all cases.</td></tr>
<tr><td>4. Searching for the Same Solution or Result: The performance of a repeated action in order to determine if the outcome of the action is identical every time.</td></tr>
<tr><td rowspan="4">Type III: Extending</td><td>1. Expanding the range of Applicability: The application of a phenomenon to a larger range of cases than that from which it originated.</td></tr>
<tr><td>2. Removing Particulars: The removal of some contextual details in order to develop a global case.</td></tr>
<tr><td>3. Operating: The act of operating upon an object in order to generate new cases.</td></tr>
<tr><td>4. Continuing: The act of repeating an existing pattern in order to generate new cases.</td></tr>
</table>

Figure 1. Generalizing actions for domain and range. Adapted from Ellis (2007).

<table>
<tr><th colspan="3">REFLECTION GENERALIZATIONS</th></tr>
<tr><td rowspan="8">Type IV: Identification or Statement</td><td colspan="2">1. Continuing Phenomenon: The identification of a dynamic property extending beyond a specific instance.</td></tr>
<tr><td rowspan="3">2. Sameness: Statement of commonality or similarity.</td><td>Common Property: The identification of the property common to objects or situations.</td></tr>
<tr><td>Objects or Representations: The identification of objects as similar or identical.</td></tr>
<tr><td>Situations: The identification of situations as similar or identical.</td></tr>
<tr><td rowspan="4">3. General Principle: A statement of a general phenomenon.</td><td>Rule: The description of a general formula or fact.</td></tr>
<tr><td>Pattern: The identification of a general pattern.</td></tr>
<tr><td>Strategy or Procedure: The description of a method extending beyond a specific case.</td></tr>
<tr><td>Global Rule: The statement of the meaning of an object or idea.</td></tr>
<tr><td>Type V: Definition</td><td colspan="2">1. Class of Objects: The definition of a class of objects all satisfying a given relationship, pattern, or other phenomenon.</td></tr>
<tr><td rowspan="2">Type VI: Influenc</td><td colspan="2">1. Prior Idea or Strategy: The implementation of a previously-developed generalization.</td></tr>
<tr><td colspan="2">2. Modified Idea or Strategy: The adaptation of an existing generalization to apply to a new problem or situation.</td></tr>
</table>

Figure 2. Ellis' (2007) reflection generalizations

Data Collection Methods

We interviewed 20 students enrolled in multivariable calculus at a mid-size university in the northwestern U.S. The students were volunteers selected from all the multivariable calculus students enrolled during that term, and were compensated for their participation. The course topics included vectors, vector functions, curves in two and three dimensions, surfaces, partial derivatives, gradients, directional derivatives, and multiple integrals in different coordinate systems. Each student participated in a semi-structured interview that lasted about an hour. We recorded audio and written work from each of the interviews using a LiveScribe Echo Pen, which provides a recording consisting of synced audio and written work. These recordings also allowed us to create dynamic playbacks of the interviews during data analysis. The tasks and rationale for their inclusion are shown in Table 1.

Table 1. Interview tasks and rationale

Task	Rationale
1. What does domain mean? What does range mean?	The purpose of this question was to elicit how students thought about domain and range, and what they associated with the terms, when they were not tied to a specific problem or function.
2. What are the domain and range of $f(x)=4+1/(x-3)$?	This question was included to gain insight into how students operationalized their definitions for domain and range as they worked with a single-variable function.
3. What are the domain and range of $f(x,y)=x^2+y^2$?	This question was included to gain insight into how students thought about domain and range for a multivariable function. We used this task as one way to investigate how students generalized their meanings for domain and range.
4. What are the domain and range of $x^2+y^2+z^2=9$?	This question was included to gain insight into how students thought about domain and range for a multivariable function, and how they thought of domain and range for a function written in a different form than *f*(x,y).

We had two research foci and thus preformed two separate data analyses. We used a constant comparative analysis (Corbin, 2008) to answer our first two questions, (what are students' meanings for domain and range of single- and multivariable functions?) then did a second analysis using Ellis' (2007) generalizations framework to answer the second (how do students generalize those meanings?). We present the analysis methods, findings, and discussion of students' meanings for domain and range first, then we turn to the analysis, results, and discussion about students' generalizations.

Data Analysis I: Students' Meanings for Domain and Range

We used a constant comparative analysis (Corbin, 2008) to identify what meanings students held for domain and range. Researcher 1, who had done all but two of the interviews, randomly selected half of the interview transcripts and highlighted phrases relating to how students thought about domain and range. Students used words like input, output, result, function as a whole, independent variable, dependent variable, domain goes with x (or x and y), range goes with y (or z), domain goes with the horizontal axis (or plane), range goes with the vertical axis, codomain, and so on. Researcher 1 then read the other half of the transcripts, marking the same words and looking for any other words or phrases students used in thinking about and explaining domain and range. Researcher 1 then looked for themes in this collection of students' phrases, and found that they fit the following categories: (a) Domain and range are associated with specific variable symbols in an equation, (b) Domain and range are inputs and outputs, and (c) Domain and range relate to independent and dependent variables. Researcher 1 created coding criteria for each of these categories for both single and multivariable functions, and both researchers coded all of the data independently. The two researchers compared their results, discussed any differences, and agreed upon the set of codes shown in Table 2. They then used the data within each category to form descriptions of the *meanings* students held for domain and range. In the next section, we give examples of data for each category and describe students' meanings.

Table 2. Codes and criteria for meanings of domain and range

	Criteria	
Code	***Single-Variable***	***Multivariable***
Domain is x, Range is y	Student says that domain is the x values and range is the y values without reference to the notion of function. That is, the student does not mention input, output, independent variable, or dependent variable.	Student is answering a question about f(x,y) and gives a domain for x and a range for y. The student may talk about the f(x,y) or the z value, but still identifies domain as corresponding to x and range as corresponding to y.
Input / Output	Student talks about domain as an input, a value that goes into a function, or a value that "satisfies" the function. Student talks about range as an output value, a 'return value', or the 'result value.' There is a clear reference to the notion of function.	Student is answering a question about f(x,y). Student talks about domain as inputs and identifies that there are multiple inputs because it is a function of more than one variable. Student talks about range as the output, the result of the function, the 'function value,' or the function 'as a whole'. There is a clear reference to the notion of function.
Independence / Dependence	Student identifies that domain corresponds to the independent variable and range corresponds to the dependent variable.	Student is answering a question about f(x,y). Student identifies that domain corresponds to the independent variables and range corresponds to the dependent variable. The student may use the phrase 'determined by' rather than the terms independent / dependent (e.g., "z determined by x and y")

Results & Discussion I: Students' Meanings for Domain and Range of Single and Multivariable Functions

The three broad categories in Table 1 correspond to students' meanings for domain and range. Below, we consider each of these meanings in detail.

Domain is x, range is y

One meaning that students had for domain and range was that domain meant the possible values for x and range means the possible values for y. This meaning was based on the presence of symbols in the equation rather than a notion of function. That is, probing questions about why domain was x or range was y did not yield any underlying explanations of x and y as inputs, outputs, or independent/dependent variable.

The strongest evidence that some students think of domain and range as related to specific symbols is that many students said that the domain was x and the range was y for f(x,y). For instance, Adam and Gabe both defined domain as the possible x values and range as the possible y values for a single-variable function. For $f(x,y) = x^2 + y^2$, they said

Adam: It's a helix, or spinny spring looking thing. Domain and range, so the domain of this would be all real numbers for x values, so x can equal any number, and it changes what z equals, but even negative numbers squared equal positive z. And the range is all real numbers because there is no value of y for which the graph is undefined.

Gabe: So the domain [of $f(x,y) = x^2 + y^2$,] is all real numbers because it's a square so there's no restrictions. And it's the same thing with y, it's the same as the x^2.

Int.: What would it mean if I said 4 is in the domain?
Gabe: You'd just plug it in.
Int.: But do I have to say it for x and y? If I just say 4 is in my domain and I haven't specified if it's x or y?
Gabe: I look at the domain as just being x values.
Int.: So if I said 4, it would mean that $x = 4$ is in the domain?
Gabe: Yeah.
Int.: What if I made that same statement about the range, if I said 4 was in the range? What would I be looking at?
Gabe: The y value.

Adam talked about changes in x causing changes in z, indicating he understood there was a relationship between x and z. However, he said that the range was y. Thus 'range' seemed to be attached to a specific symbol, rather than the idea of the dependence of one value on another, as he had mentioned earlier. Gabe associated domain with only x values, and range with only y values for both $f(x)$ and $f(x,y)$. In summary, the meaning of domain and range for these students was that domain corresponds to x and range corresponds to y, whether the function was a single variable function or a multivariable one.

Input/Output

One way to think about function is as a machine that takes inputs and returns outputs. Many students thought about domain and range as related to this notion of function. To these students, domain meant the possible inputs to a function and range meant the possible outputs. For single-variable functions, students identified a singular output. For multivariable functions, students explained domain as corresponding to multiple inputs, as Jim did by identifying an x input and a y input. The input-output meaning often included a link between the inputs and outputs, such that each choice of input produced a particular output. For instance, Deb talked about 'return values for each x in the domain'.

Deb: In terms of $f(x) = y$, domain would be all the value that go into the function. The domain will be all of the values for x that return a unique, I think, value for y. The range would be all the return values for each x in the domain.
Jim: Domain is your input values, otherwise known as your x values. It could also represent your independent values. The range is your output, your dependent variables, y values.
[Q3] There would be two different domains. You have your x input and your y input. Your x domain and your y domain give you a range of a different variable. It's the range of z or $f(x,y)$.

Note that Jim talked about both inputs/outputs and independence and dependence, so his answer was coded as belonging to both categories. It was fairly common for students to understand domain and range in terms of both input and output and independence and dependence.

Independence / Dependence

A function may be thought of as a relationship in which the value of one variable depends on the value of another variable. For students who thought about function this way, domain meant the possible values of the independent variable and range meant the possible values of the dependent variable. Kathy gave a good example of this with equations, and Leah talked about independence and dependence graphically by thinking about a "y plane" as determined by x values. Both Leah and Phillip identified that a multivariable function has multiple independent variables.

Kathy: Domain is the range of x values that a function can have. And I guess x is just the independent variable. If the function were $f(y)$, the domain would be y. Range is the values that a function has for the given domain. Usually it's $f(\mathrm{x}) = y$. Then y has the range.

Leah: Domain is the range of values the dependent variable can take. No, it's the independent. It's the y plane determined by the x value, or the z determined by the x and y.

Phillip: [Q3] It's a function of two variables. X and y are both independent variables, rather than the dependent variable. You could say the domain is the independent variable and range is the dependent variable.

In summary, the meanings students held for domain and range included 'domain is x and range is y'; domain as input and range as output; and domain and range as related to independent and dependent variables. In the next section, we describe how we analyzed students' *generalizations* of these meanings from their meaning in $f(\mathrm{x})$ to their meaning in $f(\mathrm{x,y})$.

Data Analysis II: Coding Students' Generalizations

Our second analysis was to determine how students generalized their meanings for domain and range as they moved from working with $f(\mathrm{x})$ to thinking about $f(\mathrm{x,y})$. We based this analysis on Ellis' (2007) generalizations framework. The framework distinguishes between *generalizing actions*, which are "students' activity as they generalize" (Ellis, 2007, p. 198), and *reflection generalizations*, which are "final statements of generalization (verbal or written) or the use of a result of a prior generalization" (Ellis, 2007, p. 198). In the next subsections, we explain how we used this framework to code our own data.

Generalizing Action: Relating

Relating is a generalizing action in which "students form an association between two or more problems, situations, ideas, or mathematical objects. They relate by recalling a prior situation, inventing a new one, or focusing on similar properties or forms of mathematical objects" (Ellis, 2007, p. 198). We only found two instances of *relating situations*. One student who defined domain and range as relating to independent and dependent variables *connected back* to a physics lab in which an experiment had had such variables. A different student, who defined domain and range in terms of inputs and outputs, engaged in *creating new* by describing temperature in California as a function of temperature in Oregon, and explained that the temperature in Oregon would be the input.

Relating objects was far more common. We found that students related both equations and graphs or coordinate axes. For instance, both Leah and Mimi related the coordinate axes of R^2 to the coordinate axes of R^3:

Leah: Range is the y plane determined by the x value, or z determined by x and y. [*Relating objects: property*]

Phillip: Lets call z the dependent variable here and move the x and y to the other side. Now the domain is x and y. [*Relating objects: property*]

Mimi: You can't have negative z but I don't know if that's the domain or the range. I'm going to say it's the range, and treat the z axis like the y axis of the function. [*Relating objects: form*]

Leah and Phillip related the coordinate axes based on the property of independence and dependence, which Leah called 'determined by.' Mimi did not use a mathematical property to relate the axes, but instead seemed to see as similar the vertical position of the y axis in $f(\mathrm{x})$ and the z axis in $f(\mathrm{x,y})$.

One clear instance of relating objects by their form was the category of students who said that domain was x and range was y for both $f(\mathrm{x})$ and $f(\mathrm{x,y})$. In these cases, the presence of x and y in an equation seemed to trigger students to say that domain was the possible x values and range was the possible y values. Ian and Gabe's descriptions of domain and range are good examples:

Ian: [Q1] [Domain] is whatever the x value can be. The values the x component can be composed of. [Range] would pretty much be the same thing except for the y component.
[Q3] so whatever x is, it would be whatever values z is because that would be the radius [writes 'domain: $-z < x < z$']. And the y is the same [writes 'range: $-y < z < y$'].
[*Relating objects: form]*

Gabe: So the domain [of $f(x,y) = x^2 + y^2$,] is all real numbers because it's a square so there's no restrictions. And it's the same thing with y, it's the same as the x^2.

Int.: What would it mean if I said 4 is in the domain?

Gabe: You'd just plug it in.

Int.: But do I have to say it for x and y? If I just say 4 is in my domain and I haven't specified if it's x or y?

Gabe: I look at the domain as just being x values.

Int.: So if I said 4, it would mean that $x = 4$ is in the domain?

Gabe: Yeah.

Int.: What if I made that same statement about the range, if I said 4 was in the range? What would I be looking at?

Gabe: The y value.

For Ian and Gabe, domain meant x and range meant y. Thus what they saw as similar in f(x) and f(x,y) was that both had an x and a y. They generalized their meaning for domain and range based on the presence of the variables in the equation. This was true of all students in the 'Domain is x, Range is y' category: students who thought domain was x and range was y in both single-and multivariable functions seemed to have made that generalization based on the presence of the variables in the equations rather than based on a conceptual meaning for domain and range.

Generalizing Action: Extending

Ellis (2007) defines *extending* as a generalizing action that "involves the expansion of a pattern, relationship, or rule into a more general structure. Students who extend widen their reasoning beyond the problem, situation, or case in which it originated" (Ellis, 2007, p.198). Our students extended the range of applicability and removed particulars. The following excerpts are representative of the ways in which students engaged in extending.

Jim: [Q1] Domain is your input values, otherwise known as your x values. It could also represent your independent values. The range is your output, your dependent values, your y values.
[Q3] There would be two different domains because there are two different inputs. I guess the range could be any number just dependent on the domain, like you could put anything into the domain and you would get a range number out. Your x domain and your y domain give you a range of a different variable. So it would be, the range would be of f(x,y).
[*Extending: removing particulars*]

Bailey: I think in 2 dimensions, whatever your domain is, you put that in and that's what your output is. I suppose that's the same in 3D as well: the array of possible values I can get out of the function.
[*Extending: removing particulars*]

Deb: [Q1] The domain is all the values for *x* that return a unique value for *y*. The range would be all of the return values. In 3D, the domain is all values for *x* and *y* and the range is all values for *z*.
[Q4] I am going to use a graph because I know it's a sphere. So the domain would be all the values between… it's like R but it's kind of limited between 3 and -3 on each part. So -3 to 3 for *x*, *y*, *z*. Those are domains. The range, it won't be 3 any more because we have… I am not sure about the range. What are the return values. I'll write it as $z = \sqrt{9 - x^2 - y^2}$. Now the range would be, that is R.
[*Extending*: *expanding the range of applicability*]

What Jim saw as similar between the domain of *f*(x) and *f*(x,y) was that in each case, domain meant input. He thus extended his idea of domain-as-input to domain-as-inputs, and likewise extended the idea of 'getting a range number out' to *f*(x,y) representing that number just as *f*(x) did. We coded this as *extending: removing particulars* because Jim removed the contextual details of the problem (that is, the function $f(x,y) = x^2 + y^2$) in order to develop a global case: domain is the input(s) and range is the output. He put the actual equation while foregrounding the meaning of domain and range. Likewise, Bailey extended the idea of range being "the array of possible values I can get out of the function" to decide that range was "the same in 3D." In stating this, she removed the particulars of the specific equation as Jim had. Deb also removed particulars, extending the idea of range as a "return value" when she worked with the equation for the sphere. Deb's meaning for range in 2D had been a return value or a *z* value. However, the equation for the sphere was written differently than the other equations. Deb extended by asking herself what the return value was, then solved the equation for *z* so she could apply her meaning for range. In doing so, she extended the range of applicability because she applied a meaning to something different from which it had originated.

Reflection generalization: Identification or statement: General principle

Ellis (2007) defines a *general principle* as "a statement of a general phenomenon" (Ellis, 2007, p. 200). General principles come under the categories of 'identification or statement' in which students make their generalizations public by explicitly writing or stating them.. Our students frequently stated global rules as they tried to think about the meaning of the domain and range of *f*(x,y). That is, one way in which they made meaning of the concepts "domain of *f*(x,y)" and "range of *f*(x,y)" was to state their meaning of the concepts "domain of *f*(x)" and "range of *f*(x)", linking the meaning in each context to form a description of the general phenomenon. For example,

Mimi: Like you've got *x*, you've got *y*, and *z* is kind of like the function value. It equals *f*(x,y) kind of like $y = f(x)$. It's the dependent variable, not the independent.

Philip: The range… is the result of the function, so I guess that would be *z*. The range is … the dependent variable. *X* and *y* are both independent variables. You could give a better definition than in question 1 and say domain is the independent variable and range is the dependent variable.

Mimi and Phillip used two ideas in their meaning of range: that of the "function value" or "result of the function" and that of dependency. The function value meaning allowed Mimi to see $z = f(x,y)$ as analogous to $y = f(x)$. Likewise, Phillip saw *z* as the "result" of the function of *x* and *y.* He stated a global rule that domain corresponds to the independent variables and range corresponds to the dependent variable. In talking about the function's value or result and independence/dependence, the students were stating the *meaning* of domain and range.

Phillip's statement is a good example of the relationship between generalizing actions and reflection generalizations. Ellis (2007) notes that reflection generalizations often come on the heels of generalizing actions. Phillip *extended* his idea about “the result of a function” from the single-variable to the multivariable case, and this extension was immediately followed by a synthesizing comment about the meaning of domain and range in general.

Reflection Generalization: Influence

There are two reflection generalizations classified as Influence. The first is *prior idea or strategy*, in which a student implements a previously developed generalization. The second is a *modified idea or strategy*, in which a student adapts an existing generalization to apply to a new problem or situation. Quincy and Neil’s statements illustrate the difference well:

Quincy: [Q1] Range is how far the function spans. Range is the set of numbers the function can have.

[Q4] I think the range is 9 for this one… because that's the value on the other side of the equal sign. So it can't really range to any other values.

Neil: [Q1] Domain is the span that the x value can take on. Range is the span that the y value can take on.

[Q3] In this instance the range is z, the output value. So I would say the variables applied to the function doesn’t necessarily correspond to domain as x, range as y. So if I looked back to my definitions in question one, I could define domain and range in 3D space with domain as the span of values that can occur on the horizontal plane and I would define range to be the span of values that are dependent on the domain and span the vertical plane.

Quincy directly applied his generalization that “range is the set of numbers the function can have” to the equation for the sphere, noting that the only number the x, y, and z could add to was 9. Thus the “set” of numbers that function had consisted of one element (namely, 9). In contrast to Quincy, who implemented an existing generalization, Neil modified his existing generalization that domain was x and range was y. Since that generalization did not seem to apply to $f(x,y) = x^2 + y^2$, he adapted his idea such that to domain was the horizontal plane and range was a dependent quantity, illustrated graphically as the vertical plane.

Results & Discussion II: How Students Generalize Their Meanings for Domain and Range

We found that students generalize their meanings for domain and range by relating situations, relating objects, and extending their meanings beyond the cases in which they had originated. However, our students did not engage in all of the generalizing actions or reflection generalization that Ellis (2007) identifies. We think that this is likely an artifact of how the data were collected: Ellis’ data come from a problem-based teaching experiment focused on deriving linear relationships, while our data comes from a single interview.

Table 3. Generalizing actions for domain and range.

<table>
<tr><th colspan="3">Ellis (2007) framework</th><th>Example in domain/range data</th></tr>
<tr><td rowspan="2">Type I: Relating</td><td rowspan="2">1. Relating situations: The formation of an association between two or more problems or situations.</td><td>Connecting Back: The formation of a connection between a current situation and a previously-encountered situation.</td><td>Domain is your input values. It could also represent your independent values. I am trying to think like in terms of my physics lab where there are independent and dependent variables and you plug in the numbers that you use.</td></tr>
<tr><td>Creating New: The invention of a new situation viewed as similar to an existing situation.</td><td>Say you need to calculate temperature and you have the temperature relative to California and you have some conversion, so the input values are the temperatures in Oregon and the output values are the temperature in California.</td></tr>
</table>

<table>
<tr><td rowspan="2"></td><td rowspan="2">2. Relating objects: The formation of an association between two or more present objects.</td><td>Property: The association of objects by focusing on a property similar to both.</td><td>Lets call z the dependent variable here and move the x and y to the other side. Now the domain is x and y.</td></tr>
<tr><td>Form: The association of objects by focusing on their similar form.</td><td>You can’t have negative z but I don’t know if that’s the domain or the range. I’m going to say it’s the range, and treat the z axis like the y axis of the function.</td></tr>
<tr><td rowspan="2">Type III: Extending</td><td colspan="2">1. Expanding the range of Applicability: The application of a phenomenon to a larger range of cases than that from which it originated.</td><td>Domain is your input values, otherwise known as your x values. It could also represent your independent values. The range is your output, your dependent values, your y values.</td></tr>
<tr><td colspan="2">2. Removing Particulars: The removal of some contextual details in order to develop a global case.</td><td>I am a little fuzzy on range in 3D. I think in 2 dimensions, whatever your domain is, you put that in and that’s what your output is. I suppose that’s the same in 3D as well: the array of possible values I can get out of the function.</td></tr>
</table>

Relating objects was a common way that students generalized their meanings of domain and range. When students related equations, some saw the symbols $f(x)$ and $f(x,y)$ as serving a similar purpose in the equation, namely as the output or the “result” of the function. This allowed them to justify that range, meaning the output or result of a function, would apply to $f(x,y)$. Others related coordinate axes, some incorporating an independence/dependence meaning (e.g., Leah’s y axis determined by the x axis and z axis determined by the xy plane) and others seeming to see as similar the axes’ orientation in space (e.g., range applies to whatever axis is vertical and domain to whatever axes are horizontal). A final relation of objects was students’ seeing as similar that both $f(x)$ equations and $f(x,y)$ equations contained the same variables. Students who used this relation often said that the domain of $f(x,y)$ was x and the range was y because that was true for $f(x,y)$.

Our students also generalized by extending their meanings of domain and range in the single-variable case to the multivariable case. These extensions often involved expanding the range of applicability, such as extending the ideas of an independent x and a dependent y to an independent x and y and a dependent z or extending the idea of an input x and an output y to an input of x and y and an output z. For some students, extending involved removing particulars (like the actual equation) to focus on the meaning of domain and range (e.g., as input and output). When students extend, they place in the background the equations they are reasoning about and foreground the meaning of the concepts.

The reflection generalizations our students stated came in the form of general principles, prior ideas, and modified ideas. Ellis (2007) notes that students’ reflection generalizations often mirror their generalizing actions, and it makes sense that our students’ extensions (generalizing actions) often resulted in statements of global rules, or statements in which they used or adapted a previous generalization to incorporate the new case of multivariable functions. As with generalizing actions, not all of Ellis’ (2007) categories for reflection generalizations were present in our data. The omissions are *continuing phenomena*, *sameness*, and *definition*. The reflection generalization taxonomy for these data are in Table 4.

Table 4. Reflection generalizations for domain and range.

<table>
<tr><th colspan="3">Ellis (2007) framework</th><th>Example in domain/range data</th></tr>
<tr><td rowspan="2">Type IV: Identification or Statement</td><td rowspan="2">3. General Principle: A statement of a general phenomenon.</td><td>Rule: The description of a general formula or fact.</td><td>[Q3a] Domain of this would be all real numbers for x values, so x can equal any number, and it changes what z equals, but even negative numbers squared equal positive z. And the range is all real numbers because there is no value of y for which the graph is undefined.</td></tr>
<tr><td>Global Rule: The statement of the meaning of an object or idea.</td><td>Z is kind of like the function value. It equals $f(x,y)$ kind of like $y = f(x)$. It’s the dependent variable, not the independent.</td></tr>
<tr><td rowspan="2">Type VI: Influence</td><td colspan="2">1. Prior Idea or Strategy: The implementation of a previously-developed generalization.</td><td>[Q1] Range is the set of numbers the function can have. [Q3b] I think the range is 9 for this one… because that's the value on the other side of the equal sign. So it can't range to any other values.</td></tr>
<tr><td colspan="2">2. Modified Idea or Strategy: Z is kind of like the function value. It equals $f(x,y)$ kind of like $y = f(x)$. It’s the dependent variable, not the independent.</td><td>In this instance the range is z, the output value. So I would say the variables applied to the function doesn’t necessarily correspond to domain as x, range as y. So if I looked back to my definitions in question one, I could define domain and range in 3D space with domain as the span of values that can occur on the horizontal plane and I would define range to be the span of values that are dependent on the domain and span the vertical plane.</td></tr>
</table>

That our data contained many of Ellis’ (2007) categories for how students generalize supports the framework as useful for analyzing students’ generalizations.

Implications for Instruction

Devoting Time to Domain and Range

The actor-oriented transfer theoretical framework is useful for exploring generalization because it characterizes what students see as similar without comparing students’ perspectives to those of experts. However, judging whether students’ generalizations are congruent with experts’ ideas becomes useful when thinking about implications for instruction. For instance, some of the ways in which students related objects allowed them to generalize that the domain of $f(x,y) = z$ was x and y and the range was z. Students who formed this generalization commonly used a meaning for domain and range as input and output or independent and dependent variables along with their generalizing action of relating objects. In contrast, students who generalized incorrectly –(relating $f(x) = y$ to $f(x,y) = z$ by concluding that the x and y were present in both equations, and thus played the same role in both) seemed to not have a conceptual meaning for domain and range, but rather a definition that was a link between a word and a symbol (that is, x is domain, y is range). As it seems to be the underlying meaning the first set of students had that allowed them to relate objects in a productive way, instructors might focus on the meaning of $f(x,y)$ as a function with multiple inputs, similar to $f(x)$ (a function with one input). Many of our students stated that the interview was the first time they had thought about the domain and range of multivariable functions. Given some students' incorrect generalizations, it would likely be beneficial for instructors to devote time to talking about the domain and range of multivariable functions.

Complementing with a Focus on Covariation

We also recognize that a strong notion of input and output is not necessarily enough for students to think about function in the ways instructors intend. A generalized notion of input and output has limitations because it relies on the notion that one quantity is dependent on another. In most real world situations, the notion of independence and dependence is contrived because one quantity's value is not actually determined by another quantity's value. While it may be useful to treat one quantity as dependent for ease of calculation of simplification of some physical situation, thinking about functions in terms of covariation is crucial to students' success in calculus (Thompson & Silverman, 2008). In short, thinking covariationally means the student thinks about a function as an invariant relationship between quantities' values not necessarily coupled with a notion of input and output (Thompson, 2011). For example, consider a situation in which a person is moving and there are two quantities: the amount of distance she has traveled and the amount of time elapsed since she began traveling. One would be reasoning covariationally if a) she conceived of both quantities and their individual variation (i.e. time varies, distance varies) and b) she conceived of those quantities varying simultaneously, so that when she thinks about a person's distance traveled, she has an image of the amount of time needed to travel that distance. There is no sense of input or output *required* (though it may be present) within covariational reasoning. Inputs, outputs, independence, and dependence ideas may (i.e. elapsed time causes elapsed distance, or vice versa) arise because of the person's conception of the situation, not because one quantity has been designated as an input and one as an output. It is important to note that covariational reasoning does not preclude an approach involving input and output. Instead, it focuses on a quantitative relationship as the basis for a function from which an input-output metaphor may or may not be drawn. Thus, while this study shows ways in which one might generalize notions of input and output, it is important that multivariable functions not be presented and talked about solely in terms of input and output. While it maybe a useful way to think about domain and range, it does not guarantee that students think about functions as they need to (that is, in terms of covariation) as is useful for calculus.

Suggestions for Further Research

Our tasks included functions of one and two variables. It would be interesting to include functions of more than two variables, such as f(w,x,y,z). A task including this might yield interesting results with students who have the 'variable perspective' (i.e., domain is x and range is y) as they must now think about variables which do not appear in $f(x) = y$. That is, the symbol w does not appear in this equation and thus as students try to explain its place in f(w,x,y,z), they might reveal things about their concepts of domain and range which were not revealed in our tasks.

This study was done with multivariable calculus students, but the concepts of domain and range are used in mathematics outside of calculus. For instance, domain and range are critical in linear transformations. Thus how linear algebra students generalize ideas of domain and range would provide an additional opportunity to study generalization, as well as the meanings for domain and range students have after a higher mathematics course.

Finally, as noted earlier, domain and range were a 'case study' of generalization in higher mathematics. There are many more single- and multivariable calculus ideas in which to explore students' generalizations; of particular interest to us are how students generalize ideas of derivatives and integration.

References

Asiala, M., Cottrill, J., Dubinsky, E., & Schwingendorf, K. E. (1997). The development of students' graphical understanding of the derivative. . (1997). The development of students' graphical understanding of the derivative. *The Journal of Mathematical Behavior, 16*(4), 399-431.

Confrey, J. S., E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education, 26*(1), 66-86.

Corbin, J. S., A. L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. CA: Sage Publications

Ellis, A. B. (2007). A taxonomy for categorizing generalizations: generalizing actions and reflection generalizations. *The Journal of the Learning Sciences, 16*(2), 221-262.

Kabael, T. U. (2011). Generalizing single variable functions to two-variable functions, function machine and APOS. *Educational Sciences: Theory and Practice, 11*(1), 16.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17-20.

Martinez-Planell, R., & Trigueros, M. (2013). Graphs of functions of two variables: results from the design of instruction. *International Journal of Mathematical Education in Science and Technology, 44*, 663-672.

McCallum, W. G., Hughes-Hallet, D., Flath, D., Osgood, B.G., Gleason, A.M., Quinney, D., Frazer Lock, P., Tecosky-Feldman, J., Mumford, D., & Tucker, T.W. (2009). *Multivariable Calculus* (5th Edition ed.). Hoboken, NJ: Wiley.

Monk, G. S. (1994). Students' understanding of function in calculus courses. *The Humanistic Mathematics Network Journal, 27*, 21-27.

Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). *Foundational reasoning abilities that promote coherence in students' understandings of function* (M. P. C. C. Rasmussen Ed.). Washington, DC: Mathematical Association of America.

Orton, A. (1983). Students' understanding of differentiation. *Educational Studies in Mathematics, 14*(3), 235-250.

Rowgawski, J. (2008). *Calculus: Early Transcendentals* New York, NY: W.H. Freeman and Company.

Thomas, G. B. (2010). *Calculus: Early Transcendentals*. Boston, MA: Pearson.

Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education* (pp. 33-57). Laramie, WY: University of Wyoming.

Thompson, P.W. & Silverman, J. (2008). The concept of accumulation in calculus. In M.P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research in teaching in undergraduate mathematics* (pp.43-52). Washington, DC: Mathematical Association of America.

Trigueros, M., & Martinez-Planell, R. (2010). Geometrical representations in the learning of two-variable functions. *Educational Studies in Mathematics, 73*, 3-19.

Yerushalmy, M. (1997). Designing representations: Reasoning about functions of two variables. *Journal for Research in Mathematics Education, 28*(4), 431-466.

Zandieh, M. J. (2000). *A theoretical framework for analyzing student understanding of the concept of derivative*.

STUDENTS' CONCEPTION OF THE TEMPORAL ORDER OF DELTA AND EPSILON WITHIN THE FORMAL DEFINITION OF A LIMIT

Aditya P. Adiredja and Kendrice James
University of California, Berkeley

Studies about students' understanding of the formal definition of a limit, or the epsilon delta definition suggest that the temporal order of delta and epsilon is one of the most challenging aspects of the formal definition. While multiple studies have documented this difficulty for some students, patterns of students' reasoning about the temporal order are largely unknown. This study investigates ways that students make sense of the temporal order by focusing on the justifications students provided for their claim about the temporal order. diSessa's Knowledge in Pieces provides a suitable framework to explore the context specificity of students' knowledge as well as the potential productivity of their prior knowledge in learning.

Keywords: limit, formal definition, students' prior knowledge, Knowledge in Pieces

In February 2012, the President's Council of Advisors on Science and Technology (PCAST) called for 1 million *additional* college graduates in Science, Technology, Engineering, and Mathematics (STEM) fields based on economic forecasts (Executive Office of the President, PCAST, 2012). Within STEM, mathematics is severely underrepresented. For example, the UC Berkeley Common Data Set (University of California, Berkeley, 2011) reported that mathematics accounted for 3% of the degrees conferred, whereas engineering and the biological sciences accounted for 11% and 13% respectively.[1] Calculus is the first opportunity for students to engage with theoretical mathematics and make the transition into advanced mathematical thinking. While calculus courses often act as a gatekeeper into mathematics and other STEM majors, some exemplary mathematics programs have successfully used them as the primary source for recruiting mathematics majors (Tucker, 1996).

The formal definition of a limit of a function at a point, as given below, also known as the epsilon-delta definition, is an essential topic in mathematics majors' development that is introduced in calculus. We say that the limit of $f(x)$ as x approaches a is L, and write

$$\lim_{x \to a} f(x) = L$$

if and only if, for every number $\varepsilon > 0$, there exists a number $\delta > 0$ such that for all numbers x where $0<|x-a|<\delta$ then $|f(x)-L|<\varepsilon$.[2] The formal definition provides the technical details for how a limit works and introduces students to the rigor of calculus. Yet research shows that thoughtful efforts at instruction at most leave students – including intending and continuing mathematics majors – confused or with a procedural understanding about the definition (Cottrill et al., 1996; Oehrtman, 2008; Tall & Vinner, 1981).

---

[1] Stanford University reported similar numbers with 3.3% for mathematics and 15.1% for engineering (Stanford University, 2011).

[2] In this paper, we will often refer to the first part of the statement (for every number $\varepsilon > 0$, there exists a number $\delta > 0$), the **for-all statement**, and the later part (if $0<|x-a|<\delta$ then $|f(x)-L|<\varepsilon$), the **if-then statement** for brevity.

Although studies have sufficiently documented that the formal definition is a roadblock for most students, little is known about how students actually attempt to make sense of the topic, or about the details of their difficulties. Most studies have not prioritized students' sense making processes and the productive role of their prior knowledge (Davis & Vinner, 1986; Przenioslo, 2004; Williams, 2001). This may explain why they reported minimal success with their instructional approaches (Davis & Vinner, 1986; Tall & Vinner, 1981). Thus, understanding the difficulty in the teaching and learning of the formal definition warrants a closer look – with a focus on student cognition and with attention to students' prior knowledge. It also calls for a theoretical and analytical framework that focuses on understanding the nature and role of students' intuitive knowledge in the process of learning.

A small subset of the studies have begun exploring more specifically student understanding of the formal definition (Boester, 2008; Knapp and Oehrtman, 2005; Roh, 2009; Swinyard, 2011, Swinyard and Larsen, 2012). They suggest that students' understanding of a crucial relationship between two quantities, ε and δ within the formal definition warrants further investigation. Davis and Vinner (1986) call it the *temporal order* between ε and δ, that is the sequential ordering of ε and δ within the formal definition where ε comes first, then δ (p. 295). They found that students often neglect its important role. Swinyard (2011) found that the relationship between the two quantities is one of the most challenging aspects of the formal definition for students. Knapp and Oehrtman (2005) and Roh (2009) document this difficulty for advanced calculus students. This difficulty is also prevalent among the majority of calculus students who struggled with the formal definition in Boester (2008). How students reason about the temporal order still remains an open question.

This study is a part of a larger study investigating the role of prior knowledge in student understanding of the formal definition. It specifically explores the claim that students struggle to understand the temporal order of ε and δ within the formal definition. We aim to answer the following research questions:

1. What claims do students make about the temporal order of ε and δ?
2. How do students reason about the temporal order of ε and δ?

Theoretical Framework

The Knowledge in Pieces (KiP) theoretical framework (Campbell, 2011; diSessa, 1993; Smith et al., 1993) argues that knowledge can be modeled as a system of diverse elements and complex connections. From this perspective, uncovering the fine-grained structure of student knowledge is a major focus of investigation, and simply characterizing student knowledge as misconceptions is viewed as an uninformative endeavor (Smith et al, 1993). Knowledge elements are context-specific; the problem is often inappropriate generalization to another context (Smith et al, 1993). For example, "multiplication always makes a number bigger" is not a misconception that just needs to be removed from students' way of thinking. Although this assertion would be incorrect in the context of multiplying numbers less than 1, when applied in the context of multiplying numbers greater than 1, it would be correct. Paying attention to contexts, KiP considers this kind of intuitive knowledge a potentially productive resource in learning (Smith et al., 1993). This means that instead of focusing on efforts to replace misconceptions, KiP focuses on characterizing the knowledge elements and the mechanisms by which they are incorporated into, refined and/or elaborated to become a new conception (Smith et al., 1993). Similarly, we view students' prior knowledge as potentially productive resources for learning. We focus our investigation on students' reasoning as potentially productive resources, and we focus our attention on the context specificity of students' knowledge.

Methods

The data for this report comes from a larger study with 25 students (18 new students, and 7 students from the pilot study reported last year) investigating the role of prior knowledge in student understanding of the temporal order. Each of these students has received some form of instruction on the formal definition during their first semester calculus course. We anticipate some knowledge about the definition to be a part of their prior knowledge. Participants of the study were racially diverse (1 African American, 11 Asian, 6 Hispanic/Latino, 1 Native Hawaiian, 5 White (Non-hispanic), 1 North African) and have different majors. In the presentation of the analysis, whenever we discuss an utterance from a student, we include the student's gender and race with the student's quote in order to give a better representation of the student whose knowledge we are discussing. This also helped us to stay mindful of any other resources outside of past instruction that might be relevant (e.g. home language). Students' names are all pseudonyms.

The protocol was designed to elicit student understanding of the formal definition, but more specifically their understanding of the relationship between delta and epsilon. To explore the stability and context specificity of students' knowledge across different contexts, we asked students about the temporal order of the two variables in four different contexts: dependence, their *temporal order*, set, and lastly we asked students to order x, $f(x)$, ε and δ according to the definition. The actual interview questions are included in the appendix. Each individual interview lasted about 2 to 3 hours. These interviews were videotaped following recommendations in Derry et al. (2010).

The first part of the analysis categorized students' response to each question about the temporal order. The three categories were: epsilon first, delta first or no order. Students responded to four questions related to the temporal order. The response to each question was given a score from 0 to 2 (delta first=0, no order=1, epsilon first=2). The sum of the score ranged from 0 to 8 and their total score placed them along a continuum between the claim of delta first and epsilon first. For students from the pilot study, scoring 2 on all the questions that were asked would lead to a total score of 8. In the first round of pilot study, students were asked only one question about the temporal order (question 1, above). In the second round of pilot study, students were asked three of the four questions (questions 1, 2 and 4). In those cases, the total would be normalized to 8 based on the number of available questions.

The second part of the analysis identified reasoning patterns from students' justifications for the temporal order. To identify reasoning patterns, we started by recording students' justification for each temporal order question. A justification included details about what the student attended to and the meaning they attached to it. We first sorted justifications according to the temporal order they supported: epsilon first, no order or delta first. At times a student started with one claim for the temporal order, but changed their mind afterwards. In this case, the justification for each claim was recorded as two different justifications and was sorted accordingly. Some students provided contradicting justifications to support the claim that there was no order. In this case, we would treat the two justifications as one reasoning pattern. The catalogue of reasoning patterns was developed through an iterative process of open coding (Glaser & Strauss, 1967). In documenting reasoning patterns, it was important to not infer the origin of the justification. We relied as much as possible on the particular thing the student said and attended to. For example, if a student were to say that epsilon depended on delta because we use delta to find epsilon, we recorded it as a reasoning pattern without investigating where the student could have gotten that idea. The student might have gotten the idea from the if-then statement, but unless the student explicitly attended to it, we would not include it as part of the reasoning pattern. The goal of the analysis was to show the diversity in justifications for the temporal order, and not to come up with an exhaustive list of justifications for any student in calculus.

Results

Relationship Between the ε and δ

The table below shows how each student in the study answered each question about the temporal order. The table is split into two. The top half includes students from the current study and the bottom half are students from the pilot study whose results were reported last year. Red marks questions answered with delta first. Yellow marks no order. Green marks epsilon first. Blue marks questions that were not asked or were not available.

Student	Dependence	Temporal	Set	Order	Total
Chen	0	0	0	0	0
Sheila	0	0	0	0	0
Spencer	0	0	0	0	0
Veronica	0	0	0	0	0
Patricia	0	0	0	0	0
Julia	0	0	1	0	1
Aruna	0	0	1	0	1
Jane	1	0	1	0	2
Milo	0	0	2	0	2
Jose	1	2	0	0	3
Katrina	0	2	1	0	3
Simon	1	0	2	0	3
Ryan	0	2	2	0	4
Guillermo	2	0	0	2	4
Silvia	1	2	1	0	4
Bryan	0	1	2	2	5
Roberto	2	2	1	2	7
Erin	2	2	2	2	8
David	0	N/A	N/A	N/A	0
Jacob	0	N/A	N/A	N/A	0
Adriana	0	0	N/A	0	0
Sophia	0	0	N/A	0	0
Anwar	0	0	N/A	0	0
Adam	2	2	2	N/A	8
Dean	2	N/A	N/A	N/A	8

Table 1. Students' responses to each question about the temporal order sorted from lowest to highest total and separated by current study vs. pilot study

Whereas last year we found consistency of responses across the contexts, the latest iteration of the study shows that students' conception of the temporal order is context specific. Some students were consistent across all questions. But the majority of students were able to answer with epsilon first in some context, but answered delta first in others. For example, Katrina, a female Hispanic/Latina student claimed that epsilon came first by remembering the proof procedure of breaking down the epsilon inequality. However, when she was asked to order x, $f(x)$, ε and δ, she put delta first because "the definition says that if you have delta then you have epsilon."

As we reported last year, during the pilot not all of the questions were asked. Adam and Dean scored an 8 without answering the other three questions because they answered the questions that were asked normatively, and was able to explain the formal definition accurately. To assist in parsing the table above, we charted the number of questions that students answered with epsilon first (score=2).

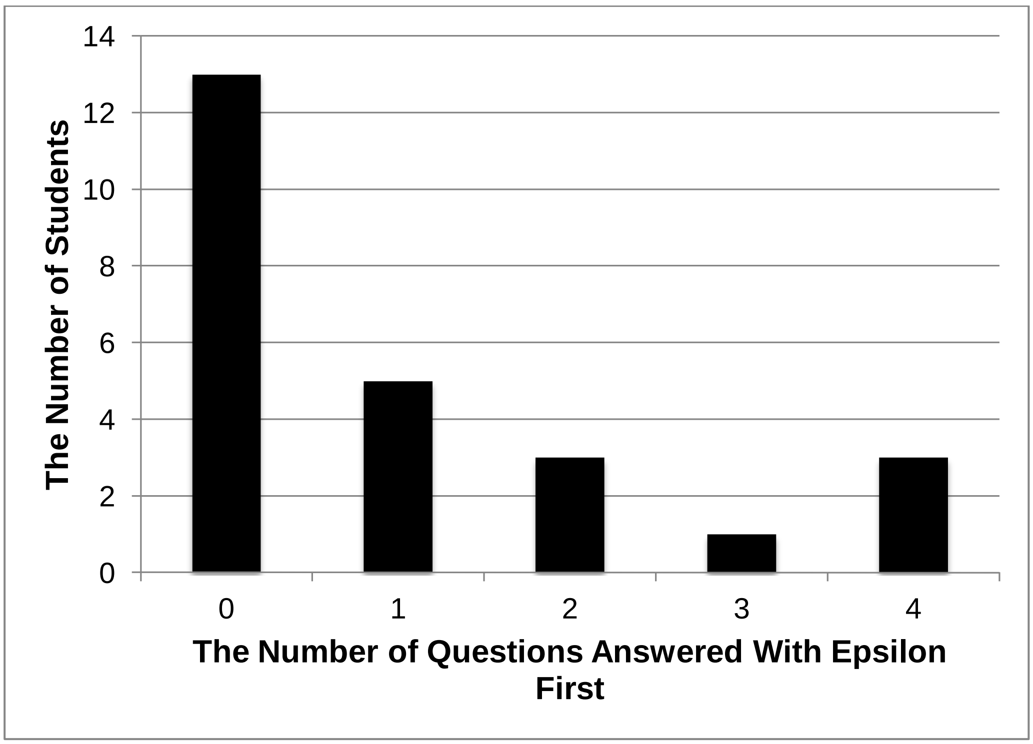


Figure 1. The distribution of students in answering the four temporal order questions with epsilon first.

Fifty two percent (52%) of students (13/25) answered none of the questions with epsilon first, while only 12% of students (3/25) answered with epsilon first on all the questions. The percentages of the rest are as follows: 1 question-20%, 2 questions-12% and 3-questions-4%. This chart shows that the majority of students in the study struggled with the temporal order of delta and epsilon.

Reasoning Patterns for the Temporal Order

The table below shows the different reasoning patterns that emerged from the data. As we said, each reasoning pattern is a type of justification students provided to support their claim about the temporal order. The table is organized by the temporal order claim for which the students used the justification. We include the number of students who used each reasoning pattern. The total number of students would exceed 25 because some students included more than one justification per question.

Temporal Order	Reasoning Pattern	Students
Delta comes first, or ε depends on δ, or δ is set first	1) Because the statement, "for every ε>0, there exists δ>0" means that there needs to be a delta (greater than zero) for the epsilon to exist.	4
	2) Because of a procedural understanding of a limit. That is, find x values close to a and check the $f(x)$ values.	6
	3) Because the if-then statement suggests that delta needs to be satisfied first then epsilon. *Note*: students might be reading the first-part of the definition, but their focus is on satisfying the delta inequality to satisfy epsilon *Variation*: the if then statement says if delta then epsilon	11
	4) Because we use delta to find epsilon.	3
	5) Because delta is related to x and epsilon is related to $f(x)$ and since $f(x)$ depends on x epsilon depends on delta. *Variation 1:* Epsilon depends on delta because $f(x)$ depends on epsilon and x depends on delta and $f(x)$ depends on x. *Variation 2:* Epsilon depends on delta because output depends on input, and delta constrains our input and epsilon constrains our output.	11
	6) Epsilon is not set because epsilon is arbitrary. So delta is set first.	4
	7) Because x and a are known, but not L. So we can use the delta inequality but not the epsilon inequality.	7
	8) Because the definition follows the order x then get delta then $f(x)$ and then epsilon. Notes: This is different from focusing on the if-then because students do not interpret the if-then question but just follow the location of each variable. *Note:* Students may look at the if-then statement but focus on the order of the quantities.	6
	9) Because of recall from the epsilon delta proof procedure, the answer is epsilon over some number.	2
	10) Because of recall from epsilon delta proof procedure, we start with the delta inequality and it will come out in the epsilon inequality.	2
No order, or ε	11) Because we have to find both of them. *Variation 1:* We are not given both epsilon and delta. *Variation 2:* If one is set, the other one is also set	7
	12) Because the for-all statement says delta depends on epsilon and the if-then statement says epsilon depends on delta	2

and δ are dependent on each other	13) Because the if-then statement suggests that they depend on each other.	3
	14) Because if the limit exists then as delta gets smaller epsilon gets smaller and if the limit doesn't exist then delta getting smaller has no effect on epsilon	1
	15) Because of proof procedure and getting a number times delta is less than epsilon	1
Epsilon comes first, or δ depends on ε, or ε is set first	16) Because the definition reads for every number epsilon, there exists a number delta, such that if $0<\|x-a\|<\delta$ then $\|f(x)-L\|<\varepsilon$, (a normative reading of the statement).	1
	17) Because of the spatial location of the variables, starting with for all.	1
	18) Epsilon is given. *Variation:* Epsilon comes first and then you find delta.	4
	19) Because of the statement for all epsilon there exists a delta	2
	20) Because of recall from the epsilon delta proof procedure, we break down the epsilon inequality to get it to look like the delta inequality or the answer is epsilon over some number.	5
	21) Since you know a and $f(x)$ you can find L, then you can set epsilon, and find delta	1
	22) Because of a counterexample where the limit does not exist and thus for a given epsilon there is no delta.	1
	23) Epsilon is set and that constrains the output which then constrains the input. *Variation*: Because we want epsilon to be really small because we want $f(x)$ to be very close to L we would want delta to be really small because we want x to be close to a.	2
	24) Because epsilon no longer depends on delta since the if then statement is about x and $f(x)$	1

Table 2. Students' reasoning patterns about the temporal order questions.

Broad Themes About the Types of Reasoning Patterns

We found quite a large number of reasoning patterns for the temporal order across the four contexts. This shows the diversity of knowledge about the temporal order. Quite a number of reasoning patterns (8/24) relied on an interpretation of different parts of the statement of the formal definition (e.g. reasoning pattern 1, 13, 16). Almost as common was those that involved a recall of the proof procedure from instruction (reasoning pattern 9, 10, 15, 20). Notice that even though students attended to the same procedure, they concluded different temporal order. Some reasoning patterns relied on a more intuitive understanding of a limit, where one would select values of x close to a to determine the limit (e.g. reasoning pattern 2, 7). Some students justified their claim using spatial location of the different variables (e.g. reasoning pattern 8, 17). So while many of these reasoning patterns might have originated from instruction, others were students' interpretation of the formal definition during the interview.

Common Reasoning Patterns in Different Contexts

One of the most common reasoning pattern, with 11 students using it as a part of their reasoning was the same justification we saw most often in our pilot study: epsilon depends on delta because delta is related to *x* and epsilon is related to $f(x)$ and since $f(x)$ depends on *x* epsilon depends on delta (reasoning pattern 5). Bryan, a White male student provided a very clear description of this reasoning pattern. He argued, "[Epsilon depends on delta] because delta is the independent variable which would be *x* in the $f(x)=y$ relationship /…/ that spits out *y*, which is our epsilon [dependent variable] /…/ because delta is like *x* and epsilon is *y*." Brian was drawing on his knowledge of functional relationships and applying that to the delta epsilon relationship. And like most students who used the functional dependence idea, Brian treated delta like *x* and epsilon like *y*.

Another reasoning pattern that was equally as common (11 students) relies on an interpretation of the if-then statement. Reasoning pattern 3 says that the if-then statement suggests that delta needs to be satisfied first before epsilon can be satisfied. So delta comes first. For example, Ryan, a male Asian student said, "For every number epsilon there is a number delta such that if the delta thing is satisfied then the epsilon is satisfied /…/ the delta has to happen for the epsilon to be satisfied. Because it goes if this, then the epsilon is satisfied. Delta needs to be satisfied before the epsilon can be." Ryan was reading the whole statement of the definition, but he clearly focused on the if-then part of the statement. He then concluded that delta came first in the temporal order. This reasoning pattern is an example of one, which relied on an interpretation of the statement. Next we explore a common reasoning pattern, which relied on students' intuitive understanding of a limit.

Seven of the 25 students argued that since the *x* and *a* were known then they could use those to find delta, whereas the limit was unknown so they could not find epsilon (reasoning pattern 7). Veronica, a White female student, argued, "Um, I would say delta [is set first] because the delta equation includes *a* whereas the components of the epsilon equation include L and you may or may not know what the limit is yet because you might be solving for the limit. But they give you *a* so I would assume that would be a better tool to use to solve." Veronica treated the delta and epsilon inequalities ($0<|x-a|<\delta$ and $|f(x)-L|<\varepsilon$) as equations. This was quite common among the students we interviewed. Doing so led her to conclude that with *x* and *a* known, she could find the delta, whereas the existence of the limit was in question. So far we have explored common reasoning patterns that support the claim that delta comes first. We explore a common one students used to argue that there was no order for epsilon and delta.

Seven students focused on the fact that they needed to find both epsilon and delta to conclude that neither was set, so there was no order (reasoning pattern 11). For example, Roberto, a male Hispanic/Latino student argued that neither epsilon nor delta were set because "you have to sort of find them or figure them out." Silvia expressed a similar opinion, "neither set because you have to solve for both of them." These students attended to whether epsilon and delta could be set, instead of which of the two was set first. We will return to his subtlety in the discussion.

It is worth noting that the most common reasoning pattern that supported students to conclude that epsilon came first was students' recall of the proof procedure (reasoning pattern 20). They were able to infer the temporal order appropriately from the proof. We will now close with one very interesting finding related to this pattern. As we mentioned earlier many students recalled voluntarily the epsilon delta proof from instruction. While five of them concluded the appropriate temporal order, many did not. In fact, many of them recalled the same procedure, attended to the same information and concluded different temporal order. For example, both Veronica (White), and Katrina (Hispanice/Latina) remembered that the "delta" would "come out" from the epsilon inequality. But Veronica concluded that delta came first, while Katrina concluded epsilon came first! Veronica said, "I'm thinking delta [comes first] because for some reason I feel like because these [$0<|x-1|<\delta$ and $3|x-1|<\varepsilon$] look kinda similar, like you can take /…/ this equation with delta and plug it in for the epsilon equation. So I'm thinking maybe you should check out delta first possibly." Katrina explained, "Oh, the one that comes first is epsilon and you figure out delta because you're gonna take this f of x minus L [$f(x)-L$] is less than epsilon and you're gonna manipulate it, and then you'll get it to look like x minus a [x-a] and depending on that, you know what delta is." The two students were both examining the inequalities and getting them to look the same, something that is commonly talked about in calculus classes, and something that they spontaneously produced during the interview. However, Veronica concluded that from this delta came first while Katrina concluded that epsilon came first. This goal of this comparison is not to compare the student's ability, but to make the point that this warrants a deeper, perhaps a more fine-grained analysis to explore what was truly underlying these conclusions and the ways in which these justifications arose.

Discussion

This study confirms the finding from our pilot study last year. We found that students struggle with the temporal order of epsilon and delta within the formal definition. Thirteen of the 25 students in this study were not able to answer *one* question about the temporal order correctly. The methods that we employed in this study allowed us to see more variability of student conceptualization of the temporal order. Ten students received a total score of 0 across the four different contexts and three students scored 8, but the majority of students were somewhere in between. The finding that some students scored a 2 in one context but a 1 or 0 in others shows that student knowledge about the temporal order is not quite stable across the different contexts. This supports our theoretical assumption that knowledge is context specific, and also highlights the importance of assessing student knowledge in multiple contexts in research and practice.

With respect to students' justifications, *the functional dependence between x and f* (x) along with *delta is with x while epsilon is with y* remain the most common reasoning pattern for the temporal order this year. We discussed the nature of that reasoning and its implication in Adiredja and James (2013). However, the current study also found another common reasoning pattern that relied on an interpretation of the if-then statement within the definition. In Adiredja and James (2013) we found that most of what we called "knowledge resources" were mathematical in nature, and hypothesized that either this indicated lack of access into the formal definition using intuitive knowledge or it was a product of using too large of a grain size to find intuitive *knowledge resources*. The findings from this study suggest that it might be both.

The findings from this study confirmed that students use their interpretation and experiences with formal mathematics to make sense of the temporal order. At the same time, a *microgenetic* case study of Adam (White male) as part of the larger project revealed that many of what we found in the pilot study were reasoning patterns, and were not quite knowledge resources. One of the authors found that a reasoning pattern is made up of different knowledge resources, making it larger in grain size. However, these reasoning patterns are useful in that it point us in a direction to locate knowledge resources. For example, the case study explore a very interesting phenomenon of a student interpreting the inequality $3|x-1|<\varepsilon$ from the proof. Sometimes Adam read that inequality to say epsilon must be greater than three times the interval around 1. Other times he read it as saying three times the interval around 1 must be smaller than epsilon. And depending on his read, he concluded the temporal order differently. The study looked into the underlying knowledge resources that influence the way he read the inequality. Perhaps the finding of that study might be informative to tease out what happened with Veronica and Katrina earlier.

We recognize one potential limitation of the current study. Four of the 19 students (Jane, Katrina, Roberto and Silvia) who were asked the set question did not interpret the question as we intended. Instead of focusing on which of the two quantities had to be set first, they were focused on whether epsilon and delta could be set. We recognize that this was a reasonable interpretation. We still coded them as no order for consistency instead of creating a new category for them. One option that we could have done but did not do was to not code their response at all, and normalize their scores much like we did with the students in the pilot study who was not asked the question. We did not do so, because we do believe that ultimately this would not dramatically change the general finding that we reported here about a lot of students struggled with the temporal order and the diversity of their reasoning patterns.

Conclusion and Implications

The list that we provided in this paper is not exhaustive, but it shows the diversity and range of student reasoning patterns. It is too early to turn our findings into some form of instructional intervention, but we believe it is important to reiterate the point we made in Adiredja and James (2013). The goal in instruction should not be to replace some of the unproductive reasoning patterns. Instead, any instructional intervention should help students reorganize these reasoning patterns while recognizing the contexts in which they might be useful (e.g. the productivity of the functional dependence relationship in multiple contexts in mathematics). More importantly, we argue that we need to get to the level of resources to truly understand *how* students reason with the temporal order, and the ideas that they prioritize. Then we can begin to think about a possible instructional approach to assist students in understanding the temporal order, and the formal definition more broadly. We would do so by honoring their prior knowledge.

References

Adiredja, A.P. & James, K. (2013). Students' knowledge resources about the temporal order of delta and epsilon. In M. Zandieh (Ed.), *Conference Papers Volumes in Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education, Denver, CO.*

Boester, T. (2008). *A design-based case study of undergraduates' conceptions of limits*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT 3314378).

Campbell, M. E. (2011). *Modeling the co-development of strategic and conceptual knowledge in mathematical problem solving.* (Doctoral dissertation). Retrieved from http://gradworks.umi.com/3498781.pdf

Cottrill, J., Dubinsky, E., Nichols, D., Schwingendorf, K., Thomas, K., Vidakovic, D. (1996). Understanding the limit concept: Beginning with a coordinated process scheme. *Journal of Mathematical Behavior, 15,* 167-192.

Davis, R. B., & Vinner, S. (1986). The notion of limit: Some seemingly unavoidable misconception stages, *Journal of Mathematical Behavior*, *5*, 281-303.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., Hall, R., Koschmann, T., Lemke, J. L., Sherin M. G. & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences, 19*(1), 3-53.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10* (2-3), 105-225.

Executive Office of the President. President's Council of Advisors on Science and Technology. *Report to the President. Engage to Excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Washington, DC: Executive Office of the President. President's Council of Advisors on Science and Technology, 2012. www.whitehouse.gov/administration/eop/ostp/pcast/docsreports (accessed September 1, 2012).

Knapp, J., & Oehrtman, M. (2005). Temporal Order in the Formal Limit Definition. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*. Melbourne: PME.

Oehrtman, M. (2008). Layers of abstraction: Theory and design for the instruction of limit concepts. In Carlson, M. P. & Rasmussen, C. (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 65-80), *MAA Notes #73.*

Przenioslo, M. (2004). Images of the limit of function formed in the course of mathematical studies at the university. *Educational Studies in Mathematics, 55*, 103–132.

Roh, K. H. (2009). An empirical study of students' understanding of a logical structure in the definition of limit via the e-strip activity. *Educational Studies in Mathematics.* DOI: 10.1007/s10649-009-9210-4.

Schoenfeld, A. H., Smith, J., & Arcavi, A. (1993). Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. IV, pp. 55-175). Hillsdale, NJ: Erlbaum.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, *3*(2), 115-163.

Stanford University, "Common Data Set: 2011-2012" (2011). *Common data set.* http://ucomm.stanford.edu/cds/2011.html (accessed September 1, 2012).

Swinyard, C. (2011). Reinventing the formal definition of limit: The case of Amy and Mike. *The Journal of Mathematical Behavior, 30*, 93-114.

Swinyard, C. & Larsen, S. (2012). Coming to understand the formal definition of limit: Insights gained from engaging students in reinvention. *Journal for Research in Mathematics Education, 43*(4), 465-493.

Tall, D.O. & Vinner, S. (1981). Concept image and concept definition in mathematics, with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*, 151-169.

Tucker, A. (1996). Models that work: Case studies in effective undergraduate mathematics program. *Notices of the AMS, 43*(11), 1356-1358.

University of California, Berkeley, "Common Data Set: 2011-2012" (2011). *Common data set.* *http://opa.berkeley.edu/statistics/cds/index.html* (accessed September 1, 2012).

Williams, S. (2001). Predications of the limit concept: An application of repertory grids. *Journal for Research in Mathematics Education, 32*, 341-367.

Appendix

Now I am going to ask you some specific questions about epsilon and delta, and after each question I am going to ask how sure you are of your answer.

1. In the definition, with epsilon and delta, what depends on what, if anything you think? Delta depends on epsilon? Epsilon depends on delta? They depend on each other? Or they do not depend on each other? And why?
Follow up: Where did you get that from? OR How does that relate to your idea that ____ depends on _____?

2. In the definition, between epsilon and delta, which one do you think comes first and which one do you figure out as a result? And why?

3. In the definition, between epsilon and delta, which one do you think is set? Epsilon? Delta? Both? Or neither? And why??

4. How would you put the four variables, epsilon, delta, x and f(x) in order in terms of which comes first in the definition? And why?
Follow up: Why did you order it that way?
Follow up: In terms of the process within the definition, how would you put the four variables in order?

STUDENT APPLICATION AND UNDERSTANDING OF THE FUNDAMENTAL THEOREM OF CALCULUS AT THE MATHEMATICS-PHYSICS INTERFACE

Rabindra R. Bajracharya[1] and John R. Thompson[1,2]
[1]Department of Physics and Astronomy and
[2]Maine Center for Research in STEM Education
University of Maine
Corresponding email: thompsonj@maine.edu

We studied students' understanding of the Fundamental Theorem of Calculus (FTC) in graphical representations that are relevant in physics contexts. Two versions of written surveys, one in mathematics and one in physics, were administered in multivariable calculus and introductory calculus-based physics classes, respectively. Individual interviews were conducted with students from the survey population. A series of FTC-based physics questions were asked during the interviews. The written and interview data have yielded evidence of several student difficulties in interpreting or applying the FTC to the problems given, including attempting to evaluate the antiderivative at individual points and using the slope rather than the area to determine the integral. The interview results further suggest that students often fail to make meaningful connections between individual elements of the FTC.

Key words: [Fundamental Theorem of Calculus, Physics, Difficulties, Problem-solving]

Introduction

We have been exploring the effect of student understanding of various concepts in mathematics on their understanding of physics concepts and vice versa. Learning physics concepts often requires the ability to interpret and manipulate the underlying mathematical representations (e.g., equations, graphs, and diagrams). A proper understanding of representations of physics concepts often requires identification of the relationship between the physics and the mathematics built into the representation as well as subsequent application of the mathematical concepts (Chi et al., 1981; Redish, 2005). Several studies in physics education research (PER) indicate connections between students' understanding of mathematics concepts and their understanding of physics concepts. Some PER findings suggest that some of the student difficulties categorized as physics difficulties may be related to the mathematics and its representations in addition to, or instead of, being difficulties with the physics (Christensen & Thompson, 2010; Christensen & Thompson, 2012; Meltzer, 2002; Pollock et al., 2007).

One interesting aspect of student understanding is the ability to relate mathematical concepts learned in a mathematics class to various physics concepts. One topic that plays a significant role in physics is the Fundamental Theorem of Calculus (FTC). The FTC is relevant in determining various physical quantities such as displacement, potential difference, work, etc. In order to fully understand the FTC, a working understanding of many concepts, such as function, rate of change, antiderivative, definite integral, etc., is needed. Research in undergraduate mathematics education attributes student difficulty with the FTC primarily to students' difficulty with the function concept (Carlson et al., 2003; Thompson, 1994; Thompson, 2008) and rates of change (Thompson, 1994).

Connecting student understanding of mathematics and physics is relevant to mathematics educators as well, since many mathematics courses use various basic physics topics for applications of mathematics concepts. In calculus, topics such as displacement, velocity and mechanical work are used as contexts for understanding integrals and derivatives. Studies have shown students using physics concepts while attempting to understand or interpret mathematical concepts (Bajracharya et al., 2012; Marrongelle, 2004). In fact, researchers have suggested the use of physical contexts (e.g., displacement, velocity, etc.) when introducing the FTC (Rosenthal, 1992; Schnepp & Nemirovsky, 2001). However, it may be that students who are unable to understand the physics concepts in the applied context may have more difficulty understanding the mathematical concepts being taught.

Similarly, physics students are often expected to be able to find connections between the rate of change (derivative) and the accumulation (definite integral) of a physical quantity (function), particularly based on graphical representations. However, to our knowledge there is no explicit research on student understanding of FTC concepts in physics, despite its ubiquitous use in various physics contexts. Researchers in physics education have studied student interpretation and use of graphs in kinematics. Beichner (1994) found that students did not recognize the physical meaning of areas under kinematics graph curves, and that students often performed slope calculations or subtracted axis values when an area calculation was required, regardless of what was graphed. We are exploring the extent to which students' understanding of the FTC affects their basic physics problem solving. Being able to distinguish whether students are struggling with the physics ideas or the underlying mathematics (or both) can inform instruction in both disciplines to help students connect the mathematics and the disciplinary contexts in which that mathematics is applied.

Theoretical Perspective

Our initial assumption about student learning was based on the constructivist perspective (Ernest, 2010). According to this perspective, students actively construct their knowledge during problem solving and reasoning using internal (e.g., concept images) and external (e.g., symbols, equations, graphs, etc.) representations. We also consider the notion that student knowledge comprises of all the information that is stored in their long-term memory, which could act as resources for executing various tasks such as problem solving and reasoning (Hammer, 2000; Redish, 2004). Depending on the way they use their mental resources, students may or may not solve a problem correctly. In particular, we are interested in probing the instances where students fail to correctly solve problems as a result of their conceptual difficulties. We have been investigating the conceptual difficulties that students have with the FTC, specifically in graphical representations, using the notion of *specific student difficulties* (Heron, 2003). According to this perspective, students manifest their difficulties through incorrect or inappropriate ideas, or flawed patterns of reasoning to specific questions. Specific difficulties are typically identified through empirical studies and are crucial for building theoretical models of student thinking because they could be used to verify those models.

The specific difficulties perspective does not necessarily speak to the origins of the difficulties being identified. There are other physics-based theoretical frameworks that address this to varying extents (e.g., misconceptions, knowledge in pieces, resources) (Chi, et al., 1981; diSessa, 1983; Hammer, 2000; Redish, 2004) Difficulties have different origins. Some can be due to the misapplication of a reasonable idea to an inappropriate context, e.g., students in a thermodynamics class sometimes treat thermodynamic work as a process-independent quantity; this is true for work done by forces associated with conservative fields, a situation commonly encountered in introductory physics. Some can be traced to an undeveloped distinction between two related concepts, e.g., the difficulty that students associate net force with velocity rather than acceleration could be accounted for by a confusion between the concepts of velocity and acceleration (Trowbridge and McDermott, 1981). Still others may be due to students believing an incorrect naïve "theory," e.g., having a conception of *impetus* in a moving body and associating that with net force (Clement, 1982).

Identification of specific student difficulties is a pragmatic approach that has led to the development of research-validated instructional strategies and materials that have improved students' conceptual understanding in many contexts across the physics curriculum (e.g., McDermott, 2001; McDermott et al. 2002).

Methodology

The data we report on here was collected through written surveys, individual clinical interviews, and mini-teaching interviews. Data were collected in second semester introductory calculus-based physics, introductory calculus, and multivariable calculus courses.

Written surveys. We constructed questions, often with parallel versions in both mathematics and physics, that either explicitly or implicitly requires the application of the FTC in a graphical interpretation. These questions were administered as written surveys in lecture sections of second-semester calculus-based introductory physics and multivariable calculus for two consecutive semesters. A total of 159 mathematics and 90 physics students participated during the first survey. During the second survey administration, 92 mathematics and 120 physics students participated. Here we focus on only one pair of questions (Fig. 1).

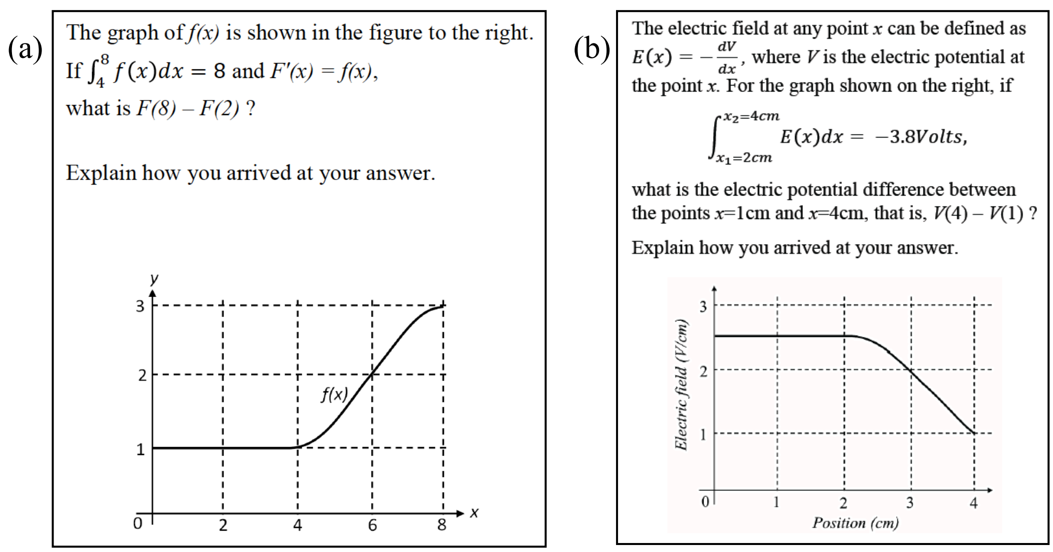


FIGURE 1. Analogous (a) mathematics and (b) physics versions of the written surveys.

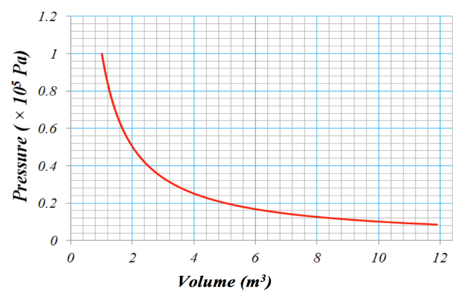
Interviews. We also conducted 14 individual interviews to probe the depth and breadth of students' understanding and application of the FTC in physics that were not revealed in the survey results, as well as the robustness of the explanations and lines of reasoning seen in the written responses. Subjects were asked four FTC-based problems in physics contexts of varying familiarity, ranging from unfamiliar to very familiar. However, these problems could be solved using the FTC without any prior knowledge of the physics. (Figure 2 depicts an example question.) The solutions to the first two problems required explicit use of the given graphs (i.e., determination of the area under the curve between the integration limits). The next two could be solved either graphically or analytically, using a given algebraic function.

Mini-teaching interviews. Once the participants solved all four interview problems, they were asked a series of calculus questions to refresh their understanding of the specific concepts. The purpose of this mini-teaching interview was to find out whether or not students could solve the problems when they were explicitly reminded of the relevant mathematics concepts required. One example of the questions asked in this part of the interview was to define, notationally, the derivative of a function *f(x)* with respect to *x*. Then they were asked to describe the process for getting the function *f(x)* back from the derivative.

1. The curve on the right shows how the pressure of a gas changes as the volume changes at constant temperature. For this curve (process), we can define pressure as the rate of change of internal energy with respect to volume, i.e.:

$$P = \frac{dU}{dV}$$

Find the change in internal energy (ΔU) when the volume changes from $V_1 = 2\times10^6$ cm^3 to $V_2 = 6\times10^6$ cm^3.



2. The figure below shows the rate of change of temperature with respect to time (i.e. *dT/dt*).
   a. Between time t = 0 and t = 1hr, how much has the temperature changed?
   b. Between t = 1hrs and t = 2hrs, how much has the temperature changed?

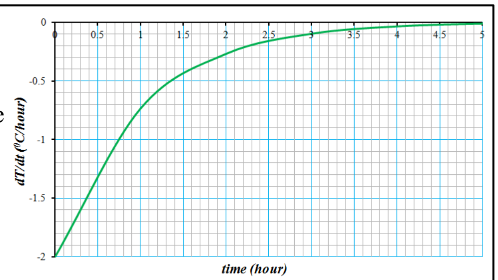


FIGURE 2. An interview problem requiring explicit use of the graph.

The initial purpose of the interviews was to probe student specific difficulties in more depth. However, as the interviews were analyzed using grounded theory (Strauss & Corbin, 1997) different kinds of interesting problem-solving strategies emerged. Thus, we also focused on the students' problem-solving strategies in our analysis in addition to any specific difficulties.

Survey Results

In the written surveys, about half of the students in both the mathematics and physics classes gave correct responses. Students used various reasoning strategies. Five strategies are described below; four of these indicate student difficulties with the FTC.

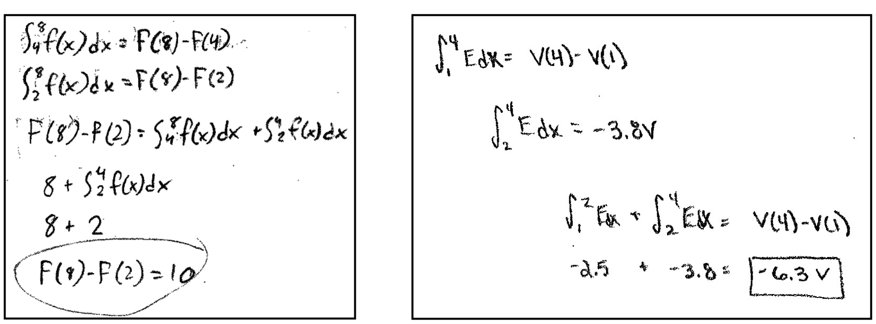


FIGURE 3. Students connecting the integral, antiderivative, and area under the curve.

1. Connecting the integral, antiderivative, and area under the curve. (Fig. 3.) Most students who provided correct responses used the FTC explicitly or implicitly. These students equated $F(b) - F(a)$, the area under the curve, and the definite integral ($\int_a^b f(x)dx$), effectively counting the squares under the curve between the limits to find the desired quantity. While most students used the correct area, a few (<5%) chose the base for their area calculations as the horizontal line that passes through the endpoint of the curve (e.g., y=1 in Fig. 1(a)) rather than the x-axis.

2. Evaluating of individual antiderivative values at endpoints. (Fig. 4.) One group of students evaluated the individual values of the antiderivatives at endpoints (e.g., $F(b)$ and $F(a)$). Finding the individual antiderivatives leads to a correct answer when they consider each of them to be equal to the areas under the curve between a common lower limit (here $F(0)$) and the upper limits as shown in Fig. 4a. However, this was not a consistently correct approach, as students also used other computational approaches to find the individual antiderivatives, as in Fig. 4b. This suggests difficulty recognizing that the difference in antiderivative values at the endpoints (e.g., $F(b) - F(a)$) is the definite integral of the given function between the given limits, and is related to the area under the curve in the given interval.

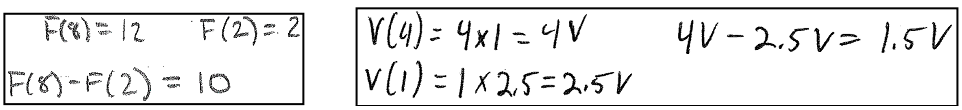


FIGURE 4. Students evaluating individual antiderivative values at endpoints.

3. Confusing antiderivative and function. (Fig. 5.) One of the most common responses was to use the difference of the original function at the endpoints (i.e. $f(b) - f(c)$) rather than the difference of the antiderivative at the endpoints (i.e. $F(b) - F(c)$), suggesting an operational confusion between the antiderivative and the function in a graphical context. This is consistent with earlier findings in upper-division thermodynamics courses in which students used the difference of endpoint values to compare the works done on a system during two different thermodynamic processes (Pollock et al., 2007).

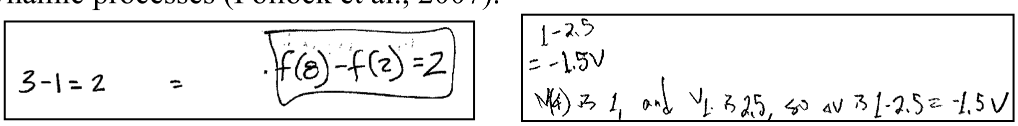


FIGURE 5. Students confusing antiderivative and function.

4. Using slope or derivative inappropriately. (Fig. 6.) A few students provided their responses using slope-based computational reasoning. Some students evaluated the slope over the interval (i.e., $\Delta y/\Delta x$) as the required answer, whereas others tried different slope-based properties, such as $F(1) = F(0)$, in their responses.

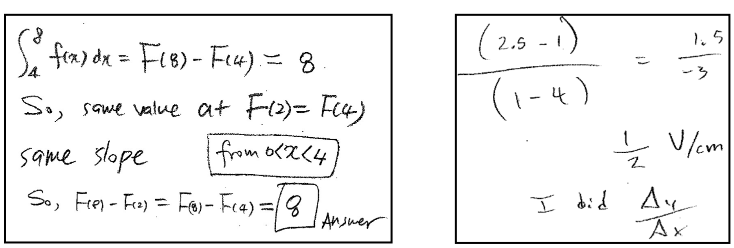


FIGURE 6. Students confusing slope or derivative with area.

5. Reasoning analytically. (Fig. 7.) Students in this category approached the problems in two distinct ways: approximating the given curve with an algebraic function, inserting that function as the integrand, and integrating; or considering the given numerical value of the integral as a function. We cannot claim to know the extent to which they understand the FTC, since their computations do not reflect relevant operations in these problems. Previous studies have also documented students' difficulties with problems without algebraic functions (Selden et al., 1989; Selden et al., 2000). This type of solution is also consistent with Dubinsky's *action* view of function (Dubinsky & Harel, 1992; Oehrtman et al., 2008).

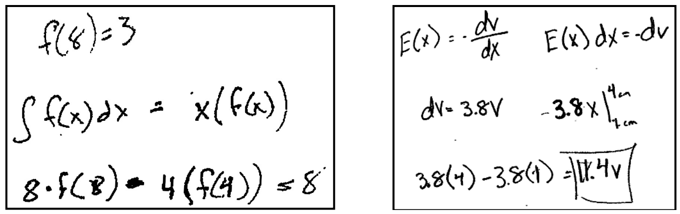


FIGURE 7. Students using analytical reasoning.

Interview Results

Our preliminary interview analysis revealed three different strategies to solve the graphically based FTC problems in physics contexts. In the first strategy, students used simple algebraic skills to rearrange the given rate equation and computational skills to produce a numerical answer. In the second strategy, students solved the problems by using graphical features, such as the area under the curve, the slope of the curve, or the difference in height of the curve at the given values. Those who used the third strategy applied their integration skills to solve the problem. In addition to integration, some of these students also used the relevant graphical feature (area under the curve), as demanded by the first two questions, whereas the others mostly did not attend to the graph. Below we illustrate three cases, each representing one type of strategy.

1. Algebraic strategy. (Fig. 8.) Although the *P–V* problem implicitly required the use of integral concepts, particularly the FTC, two-thirds of students did not use any integral concept to solve the problem. These students simply rearranged the given equation to isolate the required quantity, i.e., $dU = PdV$, and transformed it to $\Delta U = P\Delta V$ without showing any of the intermediate steps. In the following excerpt, Monica first quietly solved the *P–V* problem; when the interviewer asked her to explain the solution, she responded:

Monica: *Umm, well, I rearranged the equation so that* dU *equals the pressure times change in volume [pointing to* dU = PdV*] and then you have the change in volume and... Well, and I got change in volume [sic, energy] equals pressure times* 4×10^6 cm^3 *and then I looked at the graph for what the pressure would be at that point and then I multiplied it by approximately what the pressure looks like at that point to find* ΔU.

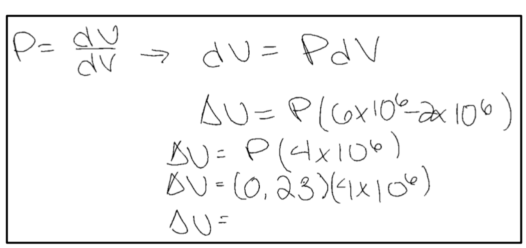


FIGURE 8. Monica's algebraic strategy to solve the *P–V* problem. (The values of V_1 and V_2 were given, whereas the value of P was extracted from the graph.)

In this example, Monica evaluated the value of ΔV using the given values of V_2 and V_1. Next, she extracted the value of P at the volume corresponding to the value of ΔV, from the graph, i.e., $P(V=4\times10^6\ \text{cm}^3)$. The required quantity (ΔU) was evaluated by taking the product of $P \cdot \Delta V$. While the specifics of Monica's strategy were unique, several students used an algebraic approach. Instead of the product of $P \cdot \Delta V$, some students found ($\Delta P \cdot \Delta V$) or ($P_1V_1 - P_2V_2$) to determine ΔU.

2. Graphical strategy. The following excerpt demonstrates how the student with pseudonym Alex responded to the $P-V$ problem. He began the solution quietly by counting the squares; when the interviewer asked him to explain what he was doing, he responded:

Alex: *I don't remember exactly how to do this problem. Umm, but, whenever there is a derivative and you are trying to find some kind of change in whether it's a pressure or depth or something like that and that's between two points... Umm, normally whenever I was in calculus, I was always taught to, basically, find the area in between. So an easy way to do that, if whenever I didn't remember how to do the problem correctly, was I always count the squares in between...*

In the above example, although Alex correctly invoked the area under the curve notion to find the change in internal energy (ΔU) between the given volumes (V_1 and V_2), he seemed to be using the notion merely as a rule – given the graph of rate of change (derivative) of a quantity, the change in quantity could be evaluated by finding the area under the rate curve. Furthermore, Alex did not use or mention anything about integration in his reasoning, further signifying his use of area under the curve merely as a rule.

Besides the area under the curve, some students also picked up on irrelevant features of the graph to solve the problem. These students determined either the average slope of the curve or the difference in height of the curve between the given values. These uses of irrelevant graphical features were also commonly seen responses to the written surveys.

3. Integral strategy. (Fig. 9.) In the following excerpt, Andrew immediately identified that the temperature problem could be solved by evaluating the area under the curve. When the interviewer asked him to explain why the area under the curve would give the required quantity, he responded by writing: $\int_0^1 dT$, then he erased dT to replace it by $\int_0^1 \frac{dT}{dt}$. After a while, he added the missing term dt to make the integral as $\int_0^1 \frac{dT}{dt}\,dt$. The interviewer then asked him:

Interview: *So how do you find the value, if I ask you to find some numerical value for change in temperature between 0 and 1?*

Andrew: *I could estimate the area under the curve, assuming that that is nearly a straight line, to make a triangle and that's square [pointing over the graph].*

In order to elicit Andrew's understanding of the connection between the integral and the area under the curve, the interviewer asked him why he thought the integral that he wrote was exactly what was being asked in the question and how the area under the curve would give the required quantity. His response was as follows:

Andrew: dT/dt *represents this curve [pointing over the graph], you integrated over time to give you change in temperature.*

Interviewer: *So what does this [showing the integral] represent in the graph?*

Andrew: *This part right here [showing the area under the curve between 0 and 1 hour].*

Interviewer: *Can you mark that with the pen?*

In response, he drew a boundary on the graph encompassing the space under the curve between 0 and 1 hour time, to show the area that represented the integral of *dT/dt*. Since there was not an equation in this problem, Andrew directly integrated the rate term (*dT/dt*) with respect to time (i.e., $\int_0^1 \frac{dT}{dt}\,dt$). However, in the $P-V$ problem, almost all the students who chose an approach similar to Andrew's first rearranged the given rate equation ($P = dU/dV$) and then integrated on both sides.

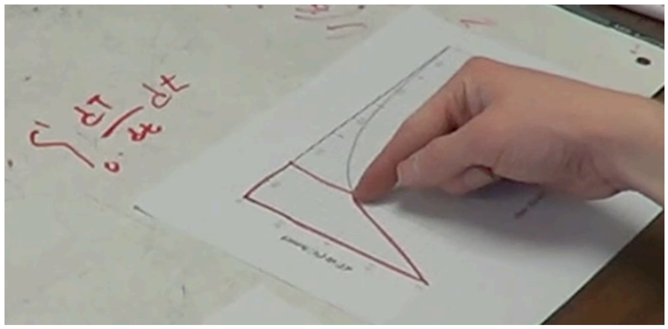


FIGURE 9. Andrew's integral-based strategy for solving the *temperature* problem. He drew the border to represent the definite integral.

Like Andrew, most of the students who correctly solved the problem using the integral approach identified that the integral could be evaluated by finding the area under the curve. Once they realized that they needed to determine the area under the curve to solve the problem, they either counted the number of squares under the curve and multiplied that number by the value of a unit square or constructed some geometric shape(s), such as right triangles and/or squares, to find the area.

Not all the students who used the integral approach evaluated the area under the curve; some determined the integral analytically. Those who chose the analytical path either approximated the given curve with an algebraic function or inappropriately considered the integrand (e.g., *P*) to be a constant to execute their integrals analytically.

Three Important Findings From The Interview Results

Besides the student problem-solving strategies discussed above, the analysis of interview data also revealed several other interesting results. These findings did not unexpectedly emerge like the student problem-solving strategies; the instruments were designed to explore the effect of the variables of interest, such as contexts, representations, knowledge elicitation, etc. Below we discuss three main findings manifested in the interviews.

1. Effects of context familiarity on student problem-solving. Since most of the interview participants were in introductory physics, they did not have any formal classroom experience on the thermodynamics context ($P-V$). The majority of students chose incorrect strategies, mostly the algebraic strategy in particular, to solve the $P-V$ problem. Some students also used the graphical strategy, i.e., area under the curve, to find the required quantity. However, most of these students did not exhibit conceptual competency in explaining why the area under the curve gives the required quantity. On the other hand, although the integrand in the temperature problem, i.e., *dT/dt*, was also not expressed in an algebraic form, a few students who used the inappropriate strategy for $P–V$ problem successfully solved the temperature problem. Although the underlying mathematical structures in both the questions were same, individual students treated the two problems quite differently indicating the attribution of the context familiarity in student problem-solving at least to some degree.

2. Effects of including integrands in algebraic forms in addition to graphical representations. The third interview problem was based in an electrostatic context, which the students had recently dealt with in their physics class. Like the first two problems, the third problem also involved a graphical representation, of electric field versus position (*E-x*). The fundamental difference between the second and the third problems was that the latter integrand was expressed as an algebraic function, whereas the former integrand was not. In the interview sample, we saw more correct solutions by a given student for problems in which both the algebraic and graphical representations were provided relative to problems with only a graphical representation. Although most of the students solved the *E-x* problem analytically in the beginning, when asked to think about alternative approach, they readily connected their analytical solution to the relevant graphical resource, i.e., the area under the curve and/or a Riemann sum representation.

Previous research on calculus concepts indicated that even those students who excel at routine problem solving often struggle with non-routine problems (Selden et al., 2000). Our results from the *P-V* problem suggest that problems involving graphical representations may be considered non-routine in this context. The *E-x* problem results further suggest that the use of algebraic representations can serve as a bridge to help students interpret graphs. Although students initially solved the *E-x* problem analytically using the integration algorithm, they seemed to be capable of dealing with the non-routine (graphical) part of the problem also once they approached it algorithmically or routinely.

3. Effects of elicitation of mathematical concepts in problem solving. In the mini-teaching interviews, almost all the students seemed have good competence with the basic calculus concepts. During this part of interview, the interviewer guided them until they arrived at the expression for the Fundamental Theorem of Calculus, and its connection with the area under the curve, either explicitly or implicitly. We found that at the end of the mini- teaching interviews, most of the participants were able to apply their mathematical knowledge to solve the physics problems. Although students initially approached the first two problems completely differently, mostly incorrectly, once their mathematical knowledge was refreshed, during the mini-teaching interviews, they were able to not only see the mathematical similarities between the two problems, but also solve them correctly. This suggests the importance of elicitation of students' mathematical knowledge relevant to the target physics contexts.

Conclusions

The preliminary results of this study describe specific student difficulties with the Fundamental Theorem of Calculus common to both mathematics and physics contexts. Some of our findings agree with previously reported difficulties, e.g., difficulties with graphical representation of integrals, relations between rate and accumulation, etc. (McDermott, et al. 1987; Beichner, 1994; Thompson, 1994). Interview results generally supported written data. The majority of students failed to use the FTC to determine the physical quantities, e.g., the change in internal energy, when the question did not include an algebraic function explicitly. For problems explicitly involving functions, most students took the antiderivative immediately and solved the problem correctly. When subsequently prompted to answer these questions using a different approach, they concluded that the solution could be represented by the area under the curve.

The interview results also revealed that the majority of the physics students – who had completed two semesters of calculus – had a reasonable grasp of most of the individual components (e.g., function, Riemann sum, definite integral, rate, etc.) of the FTC. We find that some of the specific difficulties manifested in the written surveys were the consequences of students' inabilities to access the right connections in their existing knowledge between the elements of the FTC, rather than lacking the knowledge or having a flawed understanding of these ideas, as reported previously (McDermott, et al. 1987; Beichner, 1994). Particularly, when dealing with unfamiliar physics contexts and without an analytical expression from which to start, either students struggle to meaningfully connect the individual elements of the FTC or their difficulties with even one element hinder their attempts to find these meaningful connections. We are analyzing the interviews in greater depth to see where in the protocol students recognize the appropriate connections as well as the extent to which the familiarity of the physics context affects their performance.

Our interview analyses also indicate that students use different strategies to solve the FTC-based physics problems. Although attempted initially, we did not analyze our data using the notion of transfer as preparation for future learning (e.g., Schwartz et al., 2005) because the interviews we conducted were not quite teaching interviews as suggested by Schwartz et al. Our ongoing work includes interview analysis using the lens of *epistemic games*, which are defined as a set of rules and strategies that that are guided by a specific purpose, e.g. learning a concept (Collins & Ferguson, 1993). Our approach consists of comparing our grounded-theory-based problem-solving strategies to existing, identified epistemic games to look for consistencies and inconsistencies with previous findings (Collins & Ferguson, 1993; Tuminaro 2004).

Acknowledgements

We thank the other members of University of Maine Physics Education Research Laboratory for their input and assistance with this project. We are also grateful to the course instructors for allowing us to administer written surveys and solicit students for interviews in their classes. This project was supported in part by National Science Foundation grant DUE-0817282.

References

Bajracharya, R. R., Wemyss, T. M., & Thompson, J. R. (2012). Student interpretation of the signs of definite integrals using graphical representations, *AIP Conf. Proc.* 1413, 111-114.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics, 62* (8), 750-762.

Ernest, P. (2010). Reflections on Theories of Learning, In B.Sriraman & L. English (Ed.), Theories of Mathematics Education: Seeking New Frontiers, Advances in Mathematics Education, Springer Science, Berlin/Heidelberg.

Carlson, M. P., Smith, N., & Persson, J. (2003). Developing and connecting calculus students' notation of rate of change and accumulation: The fundamental theorem of calculus. In N. A. Pateman, B. J. Doughety, & J. T. Zillioz (Ed.), *Proceedings of the 2003 Joint Meeting of PME and PMENA. 2*, pp. 165-172. Honolulu, HI: CRDG, College of Education, University of Hawai'i.

Chi, M. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5* (2), 121-152.

Christensen, W. M. and Thompson, J. R. (2012). Investigating graphical representations of slope and derivative without a physics context. *Phys. Rev. ST Physics Ed. Research* 8, 023101.

Christensen, W. M., & Thompson, J. R. (2010). Investigating student understanding of physics concepts and the underlying calculus concepts in thermodynamics. *Proceedings of the 13th Annual Conference on Research in Undergraduate Mathematics Education, MAA*.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of*

Physics, 50(1), 66-71.
Collins, A., & Ferguson, W. (1993). Epistemic forms and games. *Educational Psychologist*, 28(1), 25–42.
diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction* 10, 105-226.
Dubinsky, E., Harel, G. (1992). The nature of the process conception of function. In G. Harel & E. Dubinsky (Eds.), *The concept of functions: Aspects of epistemology and pedagogy*, MAA Notes, 25, 85-106.
Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics*, Physics Education Research Supplement, 68(S1), S52-S59.
Heron, P. (2003). Empirical investigations of learning and teaching, Part I: examining and interpreting student thinking. *Proceedings of Enrico Fermi Summer School on Physics Education Research.* (pp. 341-350). Italian Physical Society.
Marrongelle, K. A. (2004). How students use physics to reason about calculus tasks. *School Science and Mathematics*, 104(6), 258-272.
McDermott L. C. (2001). Oersted Medal Lecture 2001: Physics Education Research –The Key to Student Learning. *American Journal of Physics*. *69*, 1127-1137.
McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics, 55* (6), 503-513.
McDermott, L. C., Shaffer, P. S., and the Physics Education Group at the University of Washington (2002). *Tutorials in Introductory Physics*, 1st ed. (Prentice Hall, 2002).
Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics, 70* (12), 1259-1268.
Oehrtman, M. C., Carlson, M. P., and Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 27-42). Washington, DC: Mathematical Association of America.
Pollock, E. B., Thompson, J. R., and Mountcastle, D. B. (2007). Student understanding of the physics and mathematics of process variables in P-V diagrams. In L. Hsu (Ed.), *2006 Physics Education Research Conference* (pp. 168-171). Greensboro, NC: AIP Conference Proceedings.
Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. In *Proceedings of the Varenna Summer School*, "Enrico Fermi" Course CLVI.
Redish, E. F. (2005). Problem Solving and the Use of Math in Physics Courses. *Proceedings of the Conference: World View on Physics Education in 2005: Focusing on Change.* Delhi.
Rosenthal, B. (1992). Discovering and Experiencing the Fundamental Theorem of Calculus. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 2(2), 131-154.
Schnepp, M., & Nemirovsky, R. (2001). Constructing a Foundation for the Fundamental Theorem of Calculus. In A. A. Cuoco & F. Curcio (Eds.), *The Roles of Representation in School Mathematics* (pp. 90-102). Reston, VA: NCTM.

Schwartz D. L., Bransford J. D., Sears D. (2005). Efficiency and innovation in transfer. In: Mestre J. P., editor. *Transfer of Learning from a Modern Multidisciplinary Perspective*. Greenwich, CT: Information Age Publishing, p. 1-51.

Selden, J., Mason, A., and Selden, A. (1989). Can average calculus students solve non-routine problems? *Journal of Mathematical Behavior*, 8(1), 45-50.

Selden, A., Selden, J., Hauk, S., and Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? *Issues in Mathematics Education*, 8,128-153.

Strauss, A., & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park, AC: Sage.

Trowbridge, D. E., & McDermott, L. C. (1981). Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics*, 49(3), 242–253.

Thompson, P. W. (1994). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics, 26*, 229-274.

Thompson, P. W. (2008). The concept of accumulation in calculus. In M. P. Carlson, & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp. 43-52). Washington, DC: Mathematical Association of America.

Tuminaro, J. and Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games, *Phys. Rev. ST Phys. Educ. Res.* 3, 020101.

PROSPECTIVE SECONDARY TEACHERS' CONCEPTIONS OF PROOF AND INTERPRETATIONS OF ARGUMENTS

AnnaMarie Conner, Richard T. Francisco, Ashley L. Suominen, Carlos Nicolas Gomez, & Hyejin Park
University of Georgia

We analyzed the interviews of three prospective secondary mathematics teachers to examine their conceptions of proof and how they validated arguments in the context of students' answers. Our participants had differing views of the definition of proof and its role in mathematics, and they operationalized their conceptions of proof through differing emphases on generality, logical structure, and form or appearance of arguments. Their work when validating arguments in large part aligned with their professed views of proof, with some deviations on the part of one participant. Further research must examine whether this consistency is prevalent across prospective teachers and how this relates to teachers' work with proof in classrooms.

Key words: Proof validation, Conceptions of proof, Prospective secondary teachers, Conviction

The role of proof in mathematics has been clearly established as significant. "Proving is one of the central characteristics of mathematical behavior and probably the one that most clearly distinguishes mathematical behavior from behavior in other disciplines" (Dreyfus, 1990, p. 126). Current national recommendations establish the desirability of elementary and secondary students engaging in reasoning and proof (National Council of Teachers of Mathematics, 2009; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Teachers' conceptions of proof, beliefs about the role of proof in mathematics, and their abilities to facilitate argumentation are related to how well they can implement these kinds of experiences (see, e.g., Conner, 2007). Little research has been devoted to how prospective secondary teachers develop and modify their conceptions of proof in their university curricula. In this paper, we report results of a study in which we interviewed several prospective secondary mathematics teachers during their mathematics education coursework to examine their conceptions of proof and how they engaged in argument validation when arguments were situated in the context of student responses.

Relevant Literature

Teachers' conceptions of proof are inherently influenced by their experiences with proof in their mathematics coursework. Even though proof plays a central role in the undergraduate mathematics curriculum, numerous studies depict students' difficulties with proof production (e.g., Healy & Hoyles, 2000; Harel & Sowder, 1998). Students' lack of confidence with proof may be influenced by the fact that the field of mathematics cannot agree on a definition of proof (Hersh, 1993). However, even if students cannot give a formal definition of proof, many students have concept images of proof (Moore, 1994). Many studies have been conducted in which students at various levels were asked to construct proofs (see Reid, 2010), but as mathematics educators looked for more fine-grained explanations, some researchers have begun to examine students' validations of proofs (e.g., Knuth, 2002a; Selden & Selden, 2003; Weber, 2010).

Studies of proof validation have been conducted with various populations, including undergraduate students, practicing teachers, and research mathematicians. The results demonstrate that determining whether an argument is a valid proof is not straightforward. Selden and Selden (2003) asked undergraduate mathematics students whether given arguments proved a number theoretic statement. The aggregate of students' responses indicated a random response pattern. In another study, only six of thirteen undergraduate mathematics majors were able to determine that a real analysis proof was invalid (Weber & Alcock, 2005). A recent study on proof validation found that undergraduate students who completed an introduction to proofs course were often able to reject empirical arguments as proofs but again performed variably when asked whether a deductive argument (valid or invalid) was a proof (Weber, 2010). Research with practicing secondary teachers found that some teachers accepted non-proof arguments as valid mathematical proofs (Knuth 2002a). Finally, Weber (2008) found that even practicing mathematicians do not always agree about whether an argument is a valid mathematical proof, even for relatively uncomplicated proofs (a couple of lines long). This ambiguity has important implications for teaching, as the final verdict of a proof's correctness is often determined by social norms (e.g., Hanna, 1991). It is therefore important to examine teachers' views of proofs, what they consider to be convincing, and how they validate arguments from students.

Theoretical Perspective

Our larger study coordinates a situative perspective on learning to teach mathematics (following Peressini, Borko, Romagnano, Knuth, & Willis, 2004) with current research on teachers' beliefs about teaching, mathematics, and proof (e.g., Cooney, Shealy, & Arvold, 1998; Ernest, 1988, 1993; Knuth, 2002a; Liljedahl, Rolka, & Rosken, 2007; Thompson, 1992). As we narrowed our focus for this particular part of the study, we coordinated several perspectives related to proof to provide guidance for our analysis.

The primary lens for our analysis of participants' conceptions of proof was the multiple roles that have been proposed for proof in mathematics. Proofs provide conviction that an assertion is true (e.g., Harel & Sowder, 1998) and justify mathematical assertions. De Villiers (1990) asserted that proofs play an important communicative role in mathematics and systematize the field. Other researchers have argued that proofs should also explain why an assertion is true (e.g., Hanna, 1990; Hersh, 1993). Following from these roles of proof in the discipline of mathematics, Knuth (2002b) contended that we must consider the following roles of proof in school mathematics: verification, explanation, communication, discovery, and systematization. In Knuth's (2002b) study, practicing secondary mathematics teachers reported some of these beliefs about the role of proof, including explaining why a statement is true, communicating mathematical knowledge, verifying the truth of a statement, and systematizing the field of mathematics, but lacked emphasis on promoting understanding. In this study, we examined what our participants viewed as roles of proof in mathematics and in the classroom.

An important goal for students in teacher education programs is the development of the ability to critically reflect upon students' thinking (Ball, 1988). One prominent way that mathematical knowledge is communicated in school mathematics is through written assignments and examination. Therefore, it is essential that teachers develop proficiency at reading and analyzing mathematical arguments and proofs. We asked prospective teachers to validate mathematical arguments (after Knuth, 2002a; Selden & Selden, 2003; Weber, 2010) by stating whether they qualify as mathematical proofs and whether they find them convincing. Our analysis of our participants' argument validations was informed by Selden and Selden's (2003) description of proof validation as a process by which someone reads and reflects on an argument in order to determine the extent to which it is correct. "Validation can include asking and answering questions, assenting to claims, constructing subproofs, remembering or finding and interpreting other theorems and definitions, complying with instructions (e.g., to consider or name something), and conscious (but probably nonverbal) feelings of rightness or wrongness" (Selden & Selden, 2003, p. 5). Because we were interested in participants' views of proof in the context of teaching mathematics, we situated our interview questions and proposed arguments as answers from hypothetical secondary students. The situative perspective was useful in making sense of their responses, as they often cited norms from undergraduate mathematics classrooms or referenced classroom teaching situations when giving their evaluations.

Methodology

This paper reports a subset of results of a larger study in which we followed sixteen prospective teachers through their mathematics education coursework. For this smaller study, we purposefully selected three prospective teachers and examined their perspectives on proof during their first year of mathematics education coursework. During this time, the prospective teachers were concurrently enrolled in mathematic courses that required regular engagement with proof (e.g., Abstract Algebra and Foundation of Geometry). Data collection for this study included three video-recorded semi-structured interviews of varying length (45 – 90 minutes each). In the first interview, conducted during the first two weeks of the participants' first semester of mathematics education coursework, participants (sophomore and junior mathematics education majors) were asked for their initial thoughts on the definition of proof, its role in mathematics, and its role in the mathematics classroom. In the second and third interviews, conducted at the end of the participants' first and second semesters in the mathematics education program respectively, we asked students additional questions about proof and asked them to complete sets of proof validation tasks we had developed and adapted from other studies (see Table 1 for a summary of tasks). For example, we asked participants to examine a question or claim and then read and analyze several arguments related to the claim. For each argument, we asked the participant to state whether or not the argument was convincing and decide whether or not it proved the statement. Some of the tasks were set in the context of a classroom in which different students had proposed the different arguments. Our protocol was based in part upon Knuth's (2002b) examination of practicing teachers' beliefs about the role of proof in mathematics and in their practice, with the argument validation tasks informed by other proof validation studies as well (e.g., Weber, 2010). Each interview was transcribed by a member of the research team and checked by another member to verify accuracy.

Table 1: Summary of Arguments Presented to Prospective Teachers

Problem/Claim	Argument	Argument Summary
Exponent Problem: Is it possible to select real values for *a* and *b* such that $(2^a + 1)^b$ would be an even number? Why or why not? (Interview 2)	Cathy's	2 to any power is even so $2^a + 1$ will always be odd. An odd number to any power is odd because if "foiled" the addition of one is consistent.
	David's	2 to any power is even so $2^a + 1$ will always be odd. An odd number to any power is odd because the last digit follows a cyclic pattern of odd numbers.
The **law of cosines** states that given ΔABC with sides of length *a*, *b*, and *c* respectively, then $c^2=a^2+b^2-2ab \cos C$ (Interview 2)	A	Pre-constructed dynamic geometry sketch, steps through a series of constructions, including a circle in which are two similar triangles. A chain of equations, written from proportional relationships, concludes the argument.
	B	Dynamic geometry sketch, user can move any vertex of the triangle ΔABC and observe measurements and calculations.
	C	Two cases using Pythagorean theorem: C is

		obtuse or C is acute.
	D	Distance formula on a coordinate plane with one vertex of triangle at (0, 0).
The **sum of the first *n* odd natural numbers** is n^2. ℕ = {1, 2, 3…} (Interview 3)	Bart's	Proof by example: First 10 cases shown.
	Daphne's	Visual proof with use of multi-colored dots in square arrays.
	Charlie's	Algebraic manipulation of sum of the first *n* odd natural numbers: $S(n) = 1 + 3 + \cdots + 2n - 1$.
	Eva's	Algebraic manipulation of summation formula of the first *n* natural numbers: $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$.
	Archie's	Proof by induction.
Number Theory Problem: For any positive integers *a* and *b*, if *a* + b is an odd number, then one *a* or *b* is an odd number and the other is an even number (Interview 3)	A	Argument has unnecessary algebraic manipulation to demonstrate that an even number plus one is an odd number.
	B	Proof by contradiction. Assumes *a* and *b* are even, finding an even sum. Then assumes *a* and *b* are odd, finding an even sum.
	C	Proof of converse.
	D	Three cases: *a* is odd and *b* is even, *a* and *b* are both odd, and *a* and *b* are both even.

To analyze the data we first identified parts of the data in which participants talked about proof and proving in general, separating these from parts in which the participants were working on the proof tasks. Next, we summarized the participants' views of proof from their statements about proof and proving in general, and we summarized the participants' work with proof, paying attention to the characteristics of proof that our participants seemed to value. We coded the data on conceptions of proof and the data involving participants' validations of arguments separately. Our codes and themes were both analytic and inductive, as we began with knowledge of the purposes and characteristics of proof from the literature, but remained open to (and found) other purposes and descriptions mentioned by our participants.

Results

Our analysis of data was guided by the following questions: What are the prospective teachers' views of proof and its role in mathematics? How do the prospective teachers analyze arguments from students? What are consistencies or inconsistencies in their talk about proof and analysis of students' arguments? Our participants had differing views of the definition of proof and its role in mathematics. Their work when validating arguments in large part aligned with their professed views of proof, with some deviations on the part of one participant. In this section, we introduce Jill, Jason, and Vanessa, describe their views of proof, and briefly describe some of their argument validations.

Jill: Focus on Showing and Knowing You Are Correct

Jill focused primarily on issues of accuracy and being correct in her general talk about proving. In interview 1, she described proving as "showing that it's correct and that it works." However, she does not believe she knows "the formal definition of proving" (Interview 1), implying that there is a correct formal definition. Jill believes that we prove things in math because otherwise we would just have to take someone's word for mathematical results, so we prove things to establish mathematical certainty:

> Well if we don't prove it and somebody just says hey, this is, this works, and then they don't prove it and then how do we ever know it really does work. Because if you just, you can take anyone's word for it, but if they don't prove it and show you why it works then you might never know if it's right or not. (Interview 1)

In her examination of arguments for various statements, Jill focused on examining the details of the various steps that were given. In particular, she examined the accuracy of the algebra within three of the arguments for the law of cosines, specifically questioning how the authors obtained various lines. Likewise, she questioned a particular notation in Charlie's argument for the sum of the first n odd natural numbers, and she verified that she could see the differently sized squares in Daphne's argument for the sum of the first n odd natural numbers. Similarly, Jill paid attention to what was being proved in recognizing argument C for the number theory prompt proved the converse rather than the given statement. She also referenced specific proof techniques or notations when she was talking about her own proving as well as examining students' arguments. For instance, she stated, "Because, when we are doing, like, proofs, and we have to talk about even and odd numbers, we would usually write $2x$ for an even number and then $2x$ plus 1 for an odd, for an odd number." She was uncomfortable with Charlie's argument for the sum of the first n odd natural numbers, saying, "they just went about it in an odd way." This argument seemed to be different from what she expected, and even though she concluded that it was a proof, she seemed to be looking for a trick of some sort that would make it not a proof. In addition, Jill emphasized the need to prove all possibilities, such as in Bart's argument for the sum of odd numbers she clarifies that the student was supposed to prove if the pattern keeps going rather than working for a small set of numbers. Likewise, she said Cathy's solution to the exponent problem did not prove for all cases.

When Jill talked about proof in the context of teaching and learning, she emphasized another aspect of proof: proof as a way to understand how and why something works. In interview 1 she said proving "helps you have a reason behind things." Her explanation is similar to that of Knuth's (2002b) participants who expected their students to learn "where statements come from or why they are true rather than accepting their truth as given" (p. 80). Accordingly, Jill believes students and teachers should prove in both middle and high school when it would aid in the students' comprehension of particular concepts. Her prime example of something to be proved is the quadratic formula:

> The quadratic formula to some kids is just like a bunch of letters, and they're like, "What do I do with these letters?" I don't get it. They just plug it in and it doesn't make, they are just like, "Okay, this is what I am doing. Plug it in, blah." They don't really understand what it, what's going on, but maybe if they proved it, they would see where those letters are coming from, where the numbers go in. (Interview 3)

Jill also sees proving as useful in secondary math classes for students' understanding of a concept as well as preparing them for future mathematical ideas. She had limited proving experiences in high school; she wished she had proved more prior to college, stating, "So I think that if I had started proofs earlier or at least seen them earlier besides just geometry, then I think it would be easier and it would make more sense" (Interview 3). For her, proofs take time to understand, so students should begin proving at an earlier age than she experienced. However, despite this additional view of proof in the context of teaching and learning, when evaluating arguments, even arguments from students, Jill focused on the accuracy of the arguments, including their generality, their logical structure, and line-by-line analysis rather than the explanatory power of an argument.

When asked whether or not an argument is convincing, Jill's answers seem to be independent of whether or not she thinks the argument is a proof. For her, these are two separate questions. For instance, argument B of the law of cosines convinced her that the formula is true, especially if everything was measured correctly, but she stated that it was not a valid way to prove it. David's solution to the exponent problem was also convincing to her, but it lacked the specific proof structure or appearance to make this solution a proof. However, any argument Jill identified as a proof was also convincing to her.

Jason: An Argument is Convincing Iff it is a Proof

Jason's conceptions of proof as illustrated by his answers to general questions about proof and proving and his examination of students' arguments were very consistent. He believes that proof and proving are integral parts of mathematics. In fact, Jason brought up the idea of proof during his first interview when asked about what he thought were important aspects of math, prior to the interviewer mentioning proof. He defined proving as "demonstrating why something is the case, not just saying that's the case. So you're building up your argument" (Interview 1). Jason mentions several purposes for proof in mathematics: verification in stating, "we want to make sure we're correct" (Interview 1); explanation in stating, "it helps to know why; it helps to break it down" (Interview 1); and logic outside of mathematics in stating, "it helps think logically; you use that type of reasoning [the reasoning specifically referred to is the reasoning involved in proving that the square root of two is irrational] just all over the place" (Interview 1). In particular, he stressed the importance of proofs in relation to logic, stating "I think proving is important because then you're able to think very logically" (Interview 1) and "They [proofs] are fundamental to mathematics, understanding logic, really. It's essential to higher mathematical thinking" (Interview 2). Furthermore, in a description of a debate between graduate students discussing their relative values, Jason said he prefers a proof that is more explanatory to one that is more concise or condensed (Interview 3). This is consistent with Hersh's (1993) contention that mathematicians are more interested in why something is correct than in whether it is correct.

When he analyzed the students' arguments, he pointed out what was being proved in their arguments, investigated if the arguments included all cases and examined each step of the arguments to determine if they made sense. He distinguished between illustrating a theorem and proving it when he analyzed the dynamic geometry argument for the law of cosines (argument B), Bart's argument for the sum of the first n odd natural numbers, and Cathy's solution to the exponent problem. This is consistent with his definition of proving as demonstrating why something is the case. In several cases, Jason criticized an argument for proving something other than the requested claim, showing he paid attention to the logical structure of the arguments. This was true for two of the number theory arguments (A and C). (He critiqued both A and C as proving something not equivalent to the claim.) In his proof validations, he tended to look for generality in an argument; for instance, he critiqued Bart's argument as not proving the claim in general. He also critiqued argument B for the law of cosines: "Technically you'd have to drag the cursor over an infinite amount of screen to prove it, so no, that's not proving it" (Interview 2). Jason's view of the verification role of proof was illustrated by his answers to questions about how convincing the arguments were to him. In every case, Jason was either convinced by an argument and said it was a proof or was not convinced by an argument and said it was not a proof. This consistency was not observed in the other focus participants, and is contrary to the general trend of the findings of Segal (2000) and Weber (2010).

When asked whether or not students should participate in proving, Jason's answers were unequivocal. In the second and third interviews, he said, "Yes" (Interview 2) students should prove things, and "Absolutely" (Interview 3) students should prove things. When asked what students should prove, Jason focused more on the general concept of proving than on specific things to prove:

I think they should come across the idea of proving something is true in all cases, that just proving that something works isn't the same as proving that something is always true. I think that's an excellent concept to teach students. (Interview 2)

The specifics he mentioned regarding what students should prove, "formulas" (Interview 3) and the "quadratic formula" (Interview 2; Interview 3), seem to serve as examples, not as the totality of kinds of things students might prove. Jason also mentioned a sense-making function of proof in the classroom: "I'd say I want them to be able to come up with typical solutions that we're used to in education but also with why the solutions make sense, why the process of getting to the solutions make sense. So that's when a proof might be necessary" (Interview 3). When asked if teachers should prove, he said:

I think teachers should prove some concepts to create a model for the students to use, like 'okay this is kind of line of thought. This is how you go about proving some mathematical context,' I mean, problems. Because it'd be hard for students to come up with 'okay I'm going to come up with a counter example that is kind of clear' or contrapositive, these things need to be demonstrated. But, then once demonstrated, students can really try to do it on their own. (Interview 3)

Vanessa: Definition of Proof Depends on Audience

Vanessa's views of proof seemed to depend on her understanding of what a proof is or involves and how that coincided with the views of the instructor or the requirements of the course. Of the focus participants, Vanessa was the most accepting of arguments, including empirical arguments, as proofs. For instance, she accepted argument B for the law of cosines as a proof. In her examination of students' arguments, Vanessa did have some specific views about what a proof should look like. For instance, when examining David's solution to the exponent problem, she said that it was not what a formal proof should look like, but it made sense and was pretty convincing. She said that Archie's argument for the sum of the first n odd natural numbers was what she was used to seeing, so "I'm guessing" it's a proof (Interview 2). She critiqued Daphne's argument for the sum of the first n odd natural numbers as not a complete proof because she was used to "seeing a lot more writing and a lot more variables involved" and ultimately you cannot draw a picture for a proof (Interview 3). Vanessa also emphasized that proofs must include a generalization for all cases. In particular, she considered Cathy's argument for the exponent problem not a proof because Cathy did not generalize for all cases, whereas Vanessa was more accepting of Daphne's argument for the sum of the first n odd natural numbers since Daphne generalized the pattern of dots.

Unlike the other participants, Vanessa's definition of proof was flexible and considered the audience of the proof as an important factor, even at the beginning of her mathematics education coursework:

To prove something is when…you're able to explain the concept or an idea to someone so that they can, like, understand it. It doesn't have to be ambiguous and like just mathematically jargon-filled and, like, complicated. It can be as simple as, like, a middle school person could understand it. So it's just a way for you to be able to explain something very well, so that somebody that it's not familiar with it can be able to really understand, I think. That's when you know that you've achieved the goal of proving something. (Interview 1)

When examining Eva's argument for the sum of the first n odd natural numbers, she essentially said that it was a proof for her but not for a high school student:

But if, if I was like a high school student reading this. It wouldn't…make sense to me. It doesn't justify anything. Because I didn't know this fact [points to $1 + 2 + 3 + \ldots + n-1 + n = n(n+1)/2$], so you're telling me to assume that fact, and then once I assume it then I should believe the rest. So to a high school student this is not a proof, this doesn't explain this statement right here, this claim right here. But to me, it makes sense as proof, because I know that [points to $1 + 2 + 3 +\ldots+ n\text{-}1 + n = n(n+1)/2$], and the whole thing just follows. (Interview 3)

When she examined Bart's argument in interview 3, she distinguished that it is a justification, which is appropriate for middle school, but it is not a proof. Vanessa's flexible definition of proof could be compared to Stylianides' (2007) definition of proof in K-12 mathematics, capturing the idea of considering classroom communities, even though she does not seem to acknowledge the deductive structure implied by Stylianides.

In addition, Vanessa mentions an explanatory role for proof, both for proving in general and for why it is important in the classroom, "So I think proving, even though you understand the material, proving it makes you like know why it is true" (Interview 1) and "it's more like showing why something works, why it's true" (Interview 3). Vanessa believes that proofs can also help students to remember and retain the content. She stated, "[proof is] one of the things that help you like not just memorize a random fact but actually retain it" (Interview 1) and "once you're able to explain how you get your answer, how you got your answer, and what your thought process are, then you'll tend to remember the material that you learned better" (Interview 3). Finally, for high school students, she sees proofs as helping with making connections between concepts: "for high school I think making all that connection and seeing how everything is connected would be like the proof" (Interview 3).

Vanessa sees a continuum of proving experiences throughout middle and high school, which aligns with her classroom communities view of proof. She says:

> High school, yeah, they should have some proofs to kind of like have meaning to what they're doing. But in middle school, I wouldn't say 'proving,' I would say 'justifying.' There should be more assignments where students can't just give an answer, they should be able to justify the answer and explain the answer. And then with that they'll progress into proving their answers when they get to high school. (Interview 3)

Vanessa also sees an additional reason that teachers should prove: to demonstrate that the teachers "actually know what they're talking about" (Interview 3). However, she does not think that teachers should prove everything, "so I think teachers should prove some stuff but not prove other stuff so students get confused" (Interview 3). Likewise, teachers should not always use formal proofs: "if you're introducing new ideas or new concepts to students it's good to like maybe not a formal proof but to like explain why this formula works or why we have this" (Interview 3).

Unlike the other two participants, when asked whether or not an argument is convincing and whether or not it is a proof, Vanessa answered these two questions completely independently, i.e. a proof may or may not be convincing and a convincing argument may or may not be a proof. For instance, Vanessa believes Archie's argument for the sum of the first n odd natural numbers "might be a bit confusing for a middle school" student and thus it was not convincing despite her classifying it as a proof. However, David's argument for the exponent problem was considered to be convincing by Vanessa, but she did not consider it to be a proof.

Discussion

Although their conceptions of proof differed, we found some commonalities in our focus participants' descriptions of proofs and the roles of proof in school mathematics. However, Jason, Jill, and Vanessa operationalized these characteristics and roles differently in their validations of students' arguments. This suggests that it is important to examine multiple aspects of a teacher's conception of proof when considering how a teacher may act in the classroom; it is not enough to suggest that it is fruitful, for example, for a teacher to acknowledge an explanatory purpose for proof or to suggest that if a teacher has an explanatory view of proof, he or she will act in a particular way. Participants mentioned explanation, verification, and communication as possible role of proof, and they each mentioned an additional role that linked to life outside of the immediate course. In addition, our participants paid attention to characteristics of proof including generality, logical structure, and the form or appearance of a proof. However, their analyses of arguments led to different conclusions about whether arguments were proofs, were convincing, and were appropriate for students at different levels.

Conceptions of Roles of Proof

All of our participants mentioned an explanatory role for proofs in the context of school mathematics, as they believed that this function of proofs helps students' understanding of mathematical ideas. This finding is markedly different than Knuth's (2002b) finding in a study in which he interviewed seventeen secondary teachers: "Noticeably missing in the teachers' discussions was an explicit recognition of proof serving an explanatory capacity" (p. 80). However, Jill's focus was primarily, like some of Knuth's participants, on understanding how a formula or other result worked rather than on obtaining mathematical insight from viewing or writing a proof, even though some of her statements, such as how it would aid in the comprehension of particular concepts, could be interpreted as hinting at a more 'proof as generator of insight' function. Jason's description of the explanatory role of proof included a desire for the proof to be less concise if it explained the concepts more clearly. He seemed to be explicitly aware of a possible explanatory role of proof or that proofs could be written in ways that were or were not explanatory.

Jill and Jason both mentioned that proofs verify mathematical results, but we did not find any evidence of Vanessa conceiving of proof in this role. Verification appeared to be Jill's primary view of proof; she focused on proofs establishing the truth of a mathematical claim. Jason seemed to view verification as one of many roles of proof in mathematics: he mentioned establishing results and verifying correctness along with several other descriptors of why proof is important. In previous studies, such as Knuth (2002a, 2002b), verification has been an often-mentioned role of proof in mathematics and school mathematics.

Like some of Knuth's (2002b) participants, Jason saw a role for proof in life outside of mathematics: he wanted students to learn logic that they could apply outside of mathematics. Jill and Vanessa also mentioned wanting students to learn to prove for reasons outside of their immediate coursework. Jill wanted them to be prepared for future mathematics courses or mathematical ideas, and Vanessa wanted students to be better able to remember the content, which relates to her focus on explanation as a very important role of proof.

Finally, all three participants emphasized the communicatory role of proofs, but each emphasized a different aspect of communication. Jill emphasized communication in order to establish the truth of a statement, whereas Jason stressed the role of logic in mathematics and considered that such logical thinking supports students' reasoning in communication with others. Compared to Jill and Jason, Vanessa was much more attentive to the communication or social aspect of proofs, considering the audience of the proof as an important factor, and desired that proofs be as accessible to the intended audience as possible.

Operationalization of Conceptions in Analysis of Arguments

Although Jill, Jason, and Vanessa had similar perspectives with regard to the purposes of proof, each operationalized his or her conceptions differently. These differences came to the surface when the participants were asked to analyze student work. When we compared their validations of arguments, we found that our participants' drew on different aspects of their conceptions of proofs in their analyses of arguments. They all focused on generality, form or appearance, and logical structure to different extents in their analyses of arguments; these foci seemed to be consequences of each participant's operationalization of the roles of proof.

Jill, while paying attention to logical structure, focused more on the familiar forms and line-by-line details of arguments. Her attention to generality extended to particular notations that she associated with generating a general argument. Our interpretation of Jill's analyses of arguments shows that Jill prioritized the verification role of proof in her analysis, focusing more on whether the argument was correct and proved the correct statement than on whether the argument was explanatory or communicated to a particular audience. This was also visible in her decisions about which arguments were convincing to her. Any argument that Jill judged to be a proof she also stated was convincing; additionally, Jill was convinced by other arguments, such as argument B for the law of cosines because it verified the formula as long as everything was measured correctly. Even though she emphasized an additional aspect of proof (explanation) in the context of teaching and learning, she relied on her initial conceptions of proof (prioritizing verification) by focusing on the accuracy of students' arguments, including their generality, their logical structure, and a line-by-line analysis, rather than the explanatory power of the argument

Vanessa rarely mentioned logical structure and focused more on generality and appearance as they related to the audience of a proof. This corresponds to her emphasis on the communication and explanation roles of proof. When Vanessa analyzed Eva's argument for the sum of the first n odd numbers, she mentioned that the argument would be a proof for her, but not for middle school students or even high school students unless students were familiar with the formula for the sum of the first n natural numbers. To Vanessa, the key assumption (the formula for the sum of the first n natural numbers) in the argument was left unjustified, thereby diminishing the explanatory power of the argument. Vanessa was very consistent in her conceptions of proof and her examination of students' arguments in focusing on the audience of the proof for both cases.

Jason's view of proof as fundamental to mathematics perhaps supersedes his other views of proof in that he seemed to invoke each of his mentioned roles of proof as he analyzed arguments. His argument validations were consistent with his conceptions of proof, with different aspects of his conception of proof being apparent in his validations of different arguments. He mentioned generality, logical structure, and form or appearance as necessary in his critiques, most frequently in terms of the argument not being general, proving something different than intended (such as the converse of the statement), or not being in a form accepted by mathematicians. His analysis of Daphne's (visual rather than verbal) argument for the sum of the first n natural numbers is a case in point. He rejected this as not meeting the standards of the mathematical community (being in the wrong form), thus prioritizing the verification role of proof. However, he liked the argument as fulfilling an explanatory purpose for students. His conception of proof seems to be flexible as he called on relevant aspects of his conception of proof for critiques of the different proposed arguments.

Implications for Future Research

Even though our participants' conceptions of proofs were varied and individual, we found that their conceptions seem to have developed from their past school experiences and previous college-level mathematics courses. They applied such conceptions of proofs as valuation criteria when they evaluated students' arguments, considering issues of generality, logical structure, and form or appearance. Even though each participant mentioned several similar roles of proof, their analyses of arguments depended on how they prioritized and operationalized those roles. This implies that teachers' conceptions of proof are complex and develop over many years of different experiences. It will be difficult to conceptualize productive and unproductive conceptions of proof for teachers, yet clearly these conceptions matter at least in their analyses of students' arguments.

If we, as mathematics educators, want teachers to use proofs in ways that will promote students' understanding, we should provide opportunities for prospective teachers to consider the attributes of proofs and how they can be used to promote understanding. Our study shows that prospective secondary teachers validate students' arguments in ways that are consistent with the conceptions of proof they have developed during their school and university experiences. However, their developed conceptions seem to be individual, ranging from a flexible conception that is context-dependent and considers the audience to be an important factor, to a view that is focused on the accuracy and form of an argument, to a view that focuses on generality and the logical structure of a proof. This is true even though the participants had shared experiences in the same mathematics department and mathematics education coursework. Each of these views of proof has aspects that would be useful to teachers of secondary mathematics, but each also contains aspects that could hinder teachers' assessments of student arguments. Future research should examine if teachers' validations of students' arguments remains consistent with their views of proof when larger numbers of participants are considered. In addition, future research must examine what views of proof allow teachers to assist students in constructing and critiquing arguments in effective ways.

References

Ball, D. L. (1988). Unlearning to teach mathematics. *For the learning of mathematics, 8*(1), 40–48.

Conner, A. (2007). *Student teachers' conceptions of proof and facilitation of argumentation in secondary mathematics classrooms.* Unpublished Doctoral dissertation, The Pennsylvania State University, University Park, PA.

Cooney, T. J., Shealy, B. E., & Arvold, B (1998). Conceptualizing belief structures of preservice secondary mathematics teachers. *Journal for Research in Mathematics Education, 29*, 306–333.

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, *24*, 17–24.

Dreyfus, T. (1990). Advanced mathematical thinking. In P. Nesher, & J. Kilpatrick (Eds.), *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education* (pp.113-134). UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139013499.008

Ernest, P. (1988). The impact of beliefs on the teaching of mathematics. In P. Ernest (Ed.), *Mathematics Teaching: The State of the art* (pp. 249–254). London: Falmer. Retrieved June 12, 2008 from http://www.people.ex.ac.uk/PErnest/impact.htm.

Ernest, P. (1993). Constructivism, the psychology of learning and the nature of mathematics: Some critical issues. *Science and Education, 2*, 87–93.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6–13.

Hanna, G. (1991). Mathematical proof. In David Tall (Ed.), *Advanced mathematical thinking* (pp. 54–61). Dordrecht: Kluwer Academic Publishers.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Issues in Mathematics Education: Vol. 7. Research in collegiate mathematics education III* (pp. 234−283). Providence, RI: American Mathematical Society.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24*(4), 389–399.

Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33*(5), 379–405.

Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education, 5*(1), 61–88.

Liljedahl, P., Rolka, K., & Rosken, B. (2007). Affecting affect: The reeducation of preservice teachers' beliefs about mathematics and mathematics learning and teaching. In W. G. Martin, M. E. Strutchens, & P. C. Elliott (Eds.), *The Learning of Mathematics* (pp. 319–330). Reston: National Council of Teachers of Mathematics.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*(3), 249–266.

National Council of Teachers of Mathematics. (2009). *Focus in high school mathematics: Reasoning and sense making.* Reston, VA: Author.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics.* Washington, DC: Author. Retrieved June 1, 2012 from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Peressini, D., Borko, H., Romagnano, L., Knuth, E., & Willis, C. (2004), A conceptual framework for learning to teach secondary mathematics: A situative perspective. *Educational Studies in Mathematics, 56*(1), 67–96.

Reid, D. A. (with Knipping, C.). (2010). *Proof in mathematics education: Research, learning, and teaching.* Rotterdam, the Netherlands: Sense Publishers.
Segal, J. (2000). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior, 18,* 191–210.
Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36.
Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38,* 289–321.
Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teacher and learning* (pp. 127–146). New York: Macmillan.
Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*, 431–459.
Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning, 12*(4), 306–336.
Weber, K., & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics*, *25*(1), 34–38.

PRESENTATION OF MATRIX MULTIPLICATION
IN INTRODUCTORY LINEAR ALGEBRA TEXTBOOKS

John Paul Cook
University of Science and Arts of Oklahoma

Sepideh Stewart
University of Oklahoma

We conducted an analysis of 17 modern, introductory linear algebra textbooks to investigate presentations of matrix multiplication. Using Harel's (1987) textbook analysis framework, we examined the sequencing of matrix multiplication and its accompanying rationale. We found two principal sequences: one which first defines the operation as a linear combination of column vectors before introducing the dot product method (LC to DP), and another which invokes the dot product method before linear combinations (DP to LC). The rationale for these two trajectories varied in interesting ways. LC to DP demonstrates that solving a system of linear equations is equivalent to solving its corresponding matrix equation Ax=b. The rationale for DP to LC was less focused, opting in several cases to postpone the explanation until linear transformations are covered. We hope to initiate a discussion about the effectiveness of and pedagogical implications for these two contrasting approaches.

Key words: linear algebra, matrix multiplication, textbook analysis

Matrix multiplication is likely the first abstract multiplication that students encounter in undergraduate mathematics. It is a multiplication that does not 'multiply' in the literal sense (as with scalar multiplication or the multiplication of integers). Rather, matrix multiplication is a multiplication (in the sense of ring theory) because it is associative and distributes over matrix addition. As such, it seems reasonable to expect some hesitancy from students to accept this more abstract operation (even though the computations are relatively straightforward).

Larson and Zandieh (2013) offered three different methods students use to interpret the matrix equation Ax=b: (1) as a linear combination of the columns of A, (2) by viewing the rows of Ax as the equations in a system of linear equations, and (3) as a linear transformation acting on a vector. There are still more methods that can be used (see, for example, Carlson, 1993). Since matrix multiplication can be interpreted and defined in many different ways, how is it presented in undergraduate classrooms? How is it being explained and motivated?

While no studies were found directly examining teaching practices of matrix multiplication, a possible avenue of potential insight is to investigate presentation of matrix multiplication in linear algebra textbooks. Harel (1987) presented an analysis of linear algebra textbooks, yet our work is distinct in two important ways. First, Harel's analysis was nearly three decades ago, a significant period of time in which impactful attempts at linear algebra curriculum reform have been made (for example, Carlson, Johnson, Lay, & Porter, 1993) and an array of new textbooks have been published. Second, Harel makes no direct mention of how matrix multiplication is defined or explained. Harel's findings, however, provide a useful framework with which to conduct our analysis. He found that linear algebra textbooks varied on the basis of sequencing of content, generality of vector space models, introductory material, embodiment, and symbolization. Those tenets of Harel's framework that inform our analysis are detailed in the next section.

This paper seeks to use Harel's (1987) framework to investigate the presentation of matrix multiplication in modern, introductory linear algebra textbooks. In doing so, we sought answers to the following research questions:

- How is matrix multiplication defined in modern textbooks?
- What rationale is given for the proposed definition(s)?
- What are the pedagogical implications of any differing approaches?

Theoretical Framework

We employ Harel's (1987) framework for textbook analysis. However, as Harel's paper presented a macro-analysis (of the content presentation on a general scale throughout entire textbooks) and this paper presents a microanalysis (of the presentation of one specific topic), we adapted the framework to fit the parameters of this study. Those tenets relevant to our very specific analysis of matrix multiplication are sequencing of content and introductory material. We restrict ourselves to these two to form the basis of our analysis.

Sequencing of content

Harel noted that introductory textbooks typically follow a computation-to-abstraction approach, in which systems of equations and matrix multiplication are used to necessitate vector spaces and more general mathematical structure. Restricting our focus specifically to matrix multiplication, however, we hoped to glean insights into the overall focus and structure of the textbook by examining the sequencing of the different methods of multiplying matrices.

Introductory material

Harel found that introductory material, attempting to bridge the intellectual gap between prior knowledge and the new mathematics to be learned, was presented by means of four primary strategies:

(1) analogy: describing similarities between familiar notions and new ideas;
(2) abstraction: introducing students to specific examples before making general claims;
(3) isomorphization: presenting a familiar concept or structure that is isomorphic to the new one at hand[1];
(4) postponing: stating that the significance of a topic will be realized later when it is not currently obvious.

Indeed, matrix multiplication is an introductory topic in a first-semester linear algebra course (regardless of whether a textbook explicitly characterizes it as such). To this end, these four techniques provide an effective means with which to classify the rationale and explanations given for matrix multiplication. Any trends in this regard would provide insight not only into the overall pedagogical philosophies employed in these textbooks, but would also provide preliminary indications of how this topic is being taught in undergraduate classrooms.

Method

We narrowed our focus to introductory linear algebra textbooks (as advanced books are less likely to explicitly detail matrix multiplication) that had been published within the past

[1] Note that analogy and isomorphization seem quite similar. We shall distinguish the two by reserving isomorphization for literal cases of mathematical isomorphism; analogy is reserved for all other comparisons.

decade (as these books are more likely to be in use in undergraduate classrooms). We compiled an initial list of recently-published textbooks by (1) examining syllabi available online for introductory linear algebra courses at more than 20 large universities around the United States, (2) conducting online searches of textbook provider websites, and (3) examining the textbooks in our own respective university libraries. Overall, our list includes 17 modern, introductory linear algebra textbooks. Due to their propensity for introducing topics in very similar (if not identical) ways, textbooks sharing an author were deemed equivalent (and only counted once).

For each textbook, we examined any sections involving matrix arithmetic or matrix-vector products and also scanned the table of contents and index for any mention of these topics. Relevant pages were photocopied (or, for online books, printed out). Once all data had been collected in this manner, each textbook was analyzed using the framework detailed above. The framework then enabled us to identify trends and common themes across the entire data set.

Results

Though there are many methods that can be used to multiply two matrices, two primary methods of defining matrix multiplication emerged[2]: (1) the *linear combination of the columns method* (LC) (in which the matrix vector product Ax is defined as a linear combination of the columns of A), and (2) the vector *dot product method* (DP). Accompanying these trajectories were varying forms of explanations and rationale. Those trajectories initiating with the LC method favored isomorphization, whereas those initiating with DP were more varied. The details corresponding to these results are explicated in this section.

Sequencing of content

We examined the sequencing in which the textbook authors proceeded with the different methods of defining this operation. Two primary sequences emerged[3] that were very nearly evenly split amongst the examined textbooks. The most common method defined matrix multiplication first in terms of dot products of row and column vectors; we refer to this as the dot product method (DP) (e.g., DeFranza & Gagliardi, 2008; Poole, 2011).

Another trajectory initiated with a system of linear equations and proceeded to offer equivalent alternatives in terms of a vector equation and the matrix equation Ax=b, wherein Ax was defined (almost purely as a matter of notation) as a linear combination of the columns of A (e.g. Lay, 2011; Leon, 2010; Strang, 2009). The more general matrix product AB was then defined in terms of the matrix-vector product. We refer to this sequencing as the linear combination of the columns method (LC).

Rarely was matrix multiplication defined first in terms of linear transformations, though it did appear (e.g. Bretscher, 2012; Holt, 2012). The concept of linear transformations often appeared towards the letter half of an introductory course and thus did not play a prominent role in the definition of matrix multiplication in most texts that we examined. The following table classifies each analyzed textbook according to the sequencing of these two methods:

---

[2] There are, of course, other methods that can be used to multiply two matrices. Those listed are the most prevalent among the textbooks we examined. For additional information about the nonstandard methods, see Carlson (1993) and Strang (2009). A common method usually occurring amongst the more advanced material in these texts is to link matrix multiplication to preserving the composition of linear transformations.

[3] It is beyond the scope of this paper to detail the exact trajectory of each of the 17 textbooks. Rather, here we provide the trajectories that proved to be the most common overall.

Sequence	Textbooks Employing Specified Sequence
LC to DP (7 total)	Cheney & Kinkaid (2012); Holt (2012); Lay (2011); Leon (2010); Nicholson (2013); Spence, Insel, & Friedberg (2007); Strang (2009)
DP to LC (10 total)	Andrilli & Hecker (2009); Anton & Rorres (2010); Bretscher (2012); DeFranza & Gagliardi (2008); Kolman & Hill (2007); Larson (2012)[4]; Poole (2011); Shifrin & Adams (2010); Venit, Bishop, & Brown (2013); Williams (2012)

Textbooks invoking one approach typically followed with the other, though with less emphasis. For example, those introducing the operation with DP often used it to verify the LC method as an alternate method of calculation, referring to it when necessary (to introduce such concepts as the image of a matrix, for instance) but otherwise sparingly. On the other hand, those defining multiplication using LC tended to use the idea of linear combinations of column vectors as means with which to frame the entire text (or at least significant portions of it). In this sequencing, the DP method arose as a means to perform calculations more quickly or to calculate a single entry. To this end, the sequencing of these two methods reflected the larger focus and pedagogical strategies of each textbook.

Rationale

There were examples of each of the four categories of rationale. Typically, the rationale for the LC method invoked isomorphization, noting that defining matrix multiplication in this manner enables a system of equations to "be viewed in three different but equivalent ways: as a matrix equation, as a vector equation, or as a system of linear equations" (Lay, 2001, p. 36). The rationale for this approach, therefore, are almost purely theoretical. In contrast (and somewhat interestingly), the DP to LC method was decidedly less consistent and spanned the remaining three categories.

Analogy was usually invoked in terms of a real-world application (e.g. Larson, 2012, which uses matrices for quantities and prices at a concession stand). One particular instance of abstraction, opting for a purely mathematical motivation, framed the matrix multiplication as "a generalization of the dot product of vectors" (Andrilli & Hecker, 2009, p. 59).

Curiously, the most common method for motivating the DP method was to postpone its motivation altogether, presumably until the students have learned more about a particular topic (such as linear transformations). For example, Kolman and Hill (2007) intoned that "only a thorough understanding of the composition of functions and the relationship that exists between matrices and what are called linear transformations would show that the definition of multiplication given previously is a natural one" (p. 24). Others employing the same strategy acknowledged the fact that it is not defined component-wise, remarking that "mathematicians have introduced an alternative rule that is more useful" (Williams, 2012, p. 71) and that

[4] Larson, Edwards, & Falvo (2008) proceeded identically.

"experience has led mathematicians to the following more useful definition of matrix multiplication" (Anton & Rorres, 2010, p. 28).

Rationale	Method	Example
Analogy	DP to LC	"Another basic matrix operation is matrix multiplication. To see the usefulness of this operation, consider the following application in which matrices are helpful for organizing information. A football stadium has three concession areas …" (Larson, 2012, p. 42)
Abstraction	DP to LC	"Another useful operation is matrix multiplication, which is a generalization of the dot product of vectors." (Andrilli & Hecker, 2009, p. 59)
Isomorphization	LC to DP	"Theorem 3 provides a powerful tool for gaining insight into problems in linear algebra, because a system of linear equations may now be viewed in three different but equivalent ways: as a matrix equation, as a vector equation, or as a system of linear equations." (Lay, 2011, p. 36)
Postponing	DP to LC	"Only a thorough understanding of the composition of functions and the relationship that exists between matrices and what are called linear transformations would show that the definition of multiplication given previously is a natural one." (Kolman & Hill, 2007, p. 24)

Discussion

The sequencing in the two primary approaches to matrix multiplication can provide insight into the overarching pedagogical strategy of a textbook. The corresponding rationale provided for these two sequences predictably varies. Of particular note, however, is that the postponing technique invoked by authors using the DP approach seems to be in direct contrast to the theoretical rationale for the LC method. On one hand, the method of postponing seems to assert that "only a thorough understanding" of the composition of linear transformations could explain this definition. On the other, authors employing the LC method explain that using matrix multiplication to reframe a system of linear equations allows it to be examined from three different (yet equivalent) perspectives.

Harel (1987) asserted that "if the student does not see the rationale for a definition, the concept being defined seems very arbitrary. This has a negative motivational effect on the learning of the definition" (p. 31), suggesting that the postponing method may not be particularly effective. While a comprehensive understanding of matrix multiplication might indeed be elusive at such an introductory stage, the LC approach seems to offer an accessible, mathematically sound alternative to serve as a holdover until the complete picture can be revealed.

Additionally, Larson and Zandieh (2013) argued that viewing the matrix-vector product Ax as a linear combination of the columns of A is absolutely critical to understand span and linear independence (leading directly to vector spaces and their bases), lending credence to the value of the LC method. They also asserted that being able to move between the different interpretations is crucial to understanding the concepts at the heart of linear algebra. Though the pedagogical effectiveness of the two methods has not been directly investigated, the current literature seems to support a strong emphasis on viewing matrix multiplication in terms of linear combinations.

It is worth noting that the LC and DP approaches are not altogether incompatible. Those texts employing the LC method emphasized that the dot product definition is particularly useful for computation, and a textbook making use of the DP sequence first could very well afford ample focus to the linear combinations definition (even if it does not appear first). Thus, drawing significant conclusions from only the sequencing of matrix multiplication in an introductory textbook should be avoided. But on a more general level, the fact that the two most conspicuous methods of rationale for these respective approaches directly contradict each other suggests that this is not the case in general and that the linear combinations approach might not be adequately emphasized in many modern texts.

References

Andrilli, S. & Hecker, D. (2010). *Elementary linear algebra* (4th ed.). Massachusetts: Academic Press.

Anton, H. & Rorres, C. (2010). *Elementary linear algebra: Applications version* (10th ed.). Hoboken, NJ: Wiley.

Bretscher, O. (2012). *Linear algebra with applications* (5th ed.). New Jersey: Pearson.

Carlson, D. (1993). Teaching linear algebra: Must the fog always roll in? *The College Mathematics Journal, 24(1)*, 29-40.

Carlson, D., Johnson, C.R., Lay, D.C., & Porter, A.D. (1993). The linear algebra curriculum study group recommendations for the first course in linear algebra. *The College Mathematics Journal, 24(1)*, 41-46.

Cheney, W. & Kinkaid, D. (2012). *Linear algebra: Theory and applications* (2nd ed.). Massachusetts: Jones & Bartlett Learning.

DeFranza, J. & Gagliardi, D. (2008). *Introduction to linear algebra*. New York: McGraw-Hill.

Kolman, B. & Hill, D. (2007). *Introductory linear algebra* (9th ed.). New Jersey: Pearson.

Harel, G. (1987). Variations in linear algebra content presentation. *For the Learning of Mathematics, 7,* 29-32.

Holt, J. (2012). *Linear algebra with applications*. New York: W.H. Freeman.

Larson, R. (2012). *Elementary linear algebra* (7th ed.). Massachusetts: Houghton Mifflin.

Larson, R., Edwards, B. & Falvo, D. (2008). *Elementary linear algebra* (6th ed.). Massachusetts: Houghton Mifflin.

Lay, D. (2011). *Linear algebra and its applications* (4th ed.). New Jersey: Pearson.

Leon, S. (2010). *Linear algebra with applications* (8th ed.). New Jersey: Pearson.

Nicholson, K. (2013). *Linear algebra with applications* (7th ed.). New York: McGraw Hill.

Poole, D. (2011). *Linear algebra: A modern introduction* (3rd ed.). Massachusetts: Houghton Mifflin.

Spence, L., Insel, A. & Friedberg, S. (2007). *Elementary linear algebra: A matrix approach* (2nd ed.). New Jersey: Pearson.
Strang, G. (2009). *Introduction to linear algebra* (4th ed.). Massachusetts: Wellesley Cambridge Press.
Shifrin, T. & Adams, M. (2010). *Linear algebra: A geometric approach* (2nd ed.).. New York: W.H. Freeman.
Venit, S., Bishop, W., & Brown, J. (2013). *Elementary linear algebra* (2nd ed.). Ontario: Nelson Education.
Williams, G. (2012). *Linear algebra with applications* (8th ed.). Massachusetts: Jones & Bartlett Learning

DIFFERENTIAL PARTICIPATION IN FORMATVE ASSESSMENT AND ACHIEVEMENT IN INTRODUCTORY CALCULUS

Rebecca-Anne Dibbs University of Northern Colorado
Michael Oehrtman University of Northern Colorado

Prior formative assessment research has shown positive achievement gains when classes using formative assessment are compared to classes that do not. However, little is known about what, if any, benefits students that are not participating regularly in formative assessment gain from these assignments. The purpose of this study was to investigate the achievement of the students in two introductory calculus courses using formative assessment at the three different participation levels observed in class. Although there was no significant difference on any demographic variable other than gender and no significant difference in any achievement predictive variables between the groups of students at the different participation levels, there were significant differences in achievement on all but the first activity write-up and the final exam.

Key words: approximation framework, calculus, formative assessment

Students that leave STEM majors are most likely to do so during or immediately after completing the first semester of introductory calculus (Bressoud et al, 2013). One of the reasons participants in that study gave for leaving STEM was that they did not feel academically connected with their instructor (Bressoud et al, 2013). A possible solution to addressing this perceived disconnection is using more formative assessments in introductory calculus courses. Recent qualitative studies that flexible pedagogy and meeting students where they are at can help to build success and begin to overcome low self-efficacy (Wyatt, 2011). This psychological support is the first step to increasing the success and retention of at-risk students (Elliot & Gillen, 2013).

Formative assessments, low stakes assignments given to assess students' current level of understanding, increase student achievement (Black & Wiliam, 2009; Clark, 2011), but little is known about how implementing formative assessments facilitates this achievement gain. Regardless of the content area or age of participants, the effect size on most quantitative formative assessment studies is around 0.5 (Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, in press; Karpinski & D'Agostino, 2012). These studies show that classes where formative assessment is used do better on average on common summative assessments than those classes where no formative assessment is used; however, even in classes where formative assessment is used, not all students will regularly complete the formative assignments.

Almost all of the research on formative assessment has been quantitative quasi experimental studies (Black & McCormick, 2010; Black & Wiliam, 1998; Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, in press; Clark 2010, 2011) where a treatment class is compared to classes that do not use any formative assessment on some common summative assessment. However, there are two studies that suggest participation in formative assessment may be a predictor of student success. In the first study, low-ability math students on an aptitude pre-test that were taught using formative assessment outperformed high-ability students =who were taught with general lesson plans from the textbook on a common unit test (Chiesa & Robertson, 2000). Other studies of have found using formative to inform teaching decisions raises all students' achievement levels, though low-achieving mathematics students show the most gains in a precision taught course (Gallagher, 2006; Gallagher, Bones, & Lombe, 2006).

Although these studies are intriguing, there are several caveats. All of the research in the prior paragraph was conducted in primary schools in Great Britain and Ireland, and the mathematics content was multiplication tables, so computational fluency was the metric used to measure success. Chisea and Robertson were studying the efficacy of formative assessment for special education students, so neither the population nor the measures for success resemble undergraduate calculus students. Gallagher (2006) acknowledges these difficulties and calls for larger formative assessment studies with different student populations.

The purpose of this study was to investigate if there were achievement differences on summative assignments in a novel calculus curriculum between students completing different numbers of formative assessments during the semester. For this paper, we will distinguish between three different low participation levels: regular, sporadic, and non-participation. Students regularly participating in the formative assessments missed no more than five formative assessments during the semester; students in the sporadic participation group completed at least one but no more than six of the 12 formative assessments in the semester, while non-participants did not complete any formative assessments.

Methods

Black and Wiliam's (2009) formative assessment framework and Vygotsky's (1987) Zone of Proximal Development (ZPD) were used as the theoretical perspective of this project. There are several characterizations of the ZPD (Vygotsky, 1987); this report will focus on the scaffolding; where a learner is in their ZPD if they can complete a problem with assistance they could not complete independently. This characterization of ZPD dovetails with the second purpose of Black & Wiliam's framework (2009): formative assessment is used to engineer effective classroom discussions; where scaffolding may be given to a group of students in an efficient manner.

The study was conducted at a mid-sized doctoral granting university in the Rocky Mountain region. Sixty percent of the undergraduate population is female and 20% of the undergraduates self-identify as a member of an ethnic minority. 55% of the students at the university self-report that they are the first person in their family to attend college. Most of the students enrolled in introductory calculus major in elementary mathematics education, secondary mathematics education, mathematics, chemistry, meteorology, or geology; occasionally business majors or biology majors intending to pursue graduate work enroll in Calculus I instead of the suggested topics courses for their majors. There are a few graduate students from other disciplines enrolled in Calculus I each year to complete the admissions requirements to their programs as well. The gender distribution of Calculus I is similar to the university proportions, but there are generally fewer minority students enrolled in calculus. Approximately half of the students enrolled in Calculus I have prior experience with the course content; either by taking AP Calculus and failing to earn the credit or by failing the course at this or another post-secondary institution.

Participants were recruited from two introductory calculus courses taught using the approximation framework. This framework is built upon developing systematic reasoning about conceptually accessible approximations and error analyses but mirroring the rigorous structure of formal limit definitions and arguments (Oehrtman, 2008, 2009). This study focused on the three multi-week labs developing the most central topics in the course: limits, derivatives, and definite integrals. Each approximation lab consists of 20 questions designed to help students understand their context in terms of approximating a limit (Figure 1).

	Contextual	Graphical	Algebraic	Numerical
Unknown Value				
Approximation				
Error				
Error Bound				
Desired Accuracy				

Figure 1. Approximation framework

The limits lab is a typical example of the approximation labs. Students are given a function with a removable discontinuity that cannot be found algebraically. For the first row, students are asked to explain why algebraic techniques will not be sufficient to solve the problem, graph the function, give a variable name to the unknown *y* value of the removable discontinuity, and make a plan for approximating the *y* value of the removable discontinuity. Reading the lab and completing this unknown value row before class was the pre-lab assignment. During class, students worked in their groups to approximate the unknown y value and represent their approximation within the context, graphically, algebraically, and numerically. Most groups would end the first day of the lab working on error or error bound representations. The postlab asked students to describe what their group did in class, do a computation students who were comfortable with the first half of the lab would be able to complete, and write a short paragraph about what they did and did not understand about the lab. The next time students worked on the lab, they were expected to have completed all of the questions in their original context before class. During class, students would be regrouped so that each group member had worked on a different context the prior week – in this case every student would have had a different function. Students then presented the solution to their context to their group in class. The post-lab after class had a similar computation by was otherwise the same as the first week, and students were required to write up their original solution and compute an overestimate and underestimate within a given error bound for one of the other contexts presented during the Jigsaw on the second week.

During the week, instructors lectured over new material on Mondays and Fridays. On Tuesday, students work in groups on the approximation framework activity that week; an undergraduate teaching assistant and I help the instructor facilitate the group activities by circulating through the room, asking probing questions of students' understandings, and providing hints when groups get stuck. Students complete a formative assessment that night, and the class on Wednesday spends part of the class on discussing the formative assessment and the rest of the time covering new material. In addition to the weekly formative assessments, students complete 20 Webwork assignments throughout the 15 week semester, prepare a written report of their own answers to the approximation framework activities, and have five chapter exams and the final. Instructors meet once a week to discuss the schedule and activities for the next week. Students' individual reports of the approximation activities were group graded during the weekly coordination meeting. Instructors wrote their own unit tests, but the final exam was written and graded by all of the calculus instructors.

The courses were taught at the same time and on the same schedule by two equally experienced instructors. All of the lab questions were scored dichotomously so the inter-rater reliability of the lab write-ups was perfect, and the final exams were co-graded by the instructors. The content validity of the assessments was checked by the course coordinator and an additional expert on the approximation framework.

Before using participation level as the grouping variable in the analysis, demographic and grade predictive variables were investigated to see if there were any significant differences between participation groups. Gender, race, native language, and class were the demographic variables considered. Chi squared tests for differences were performed on all four variables. There was only one significant demographic difference between the participation levels; female students were significantly more likely to be regular participants (Table 1). Since asynchronous formative assessment, like the ones used in this study, require a greater level of organization and engagement, these assignments tend to slightly favor female students (DiPrete, 2013).

Table 1
Summary of demographic variable analysis

Demographic Variable	***p* value**
Gender (Male/Female)	.004
Race (White/Nonwhite)	.355
Native Language (English/Not English)	.651
Class (Freshman/Non-Freshman)	.802

Four variables known to predict student performance in introductory calculus were also measured: cumulative grade point average, ACT math score, Calculus Readiness Exam Score[1] , and the number of months between the end of the last math class a student took and the beginning of calculus. The final measure was self-reported, but the other three scores I obtained from the students' records. I performed ANOVA tests on each of these four quantities to see if the mean score differed across participation levels. The summary of the ANOVAs appears in Table 2. Although a Bonferroni was used correction on these and the preceding analyses, none of the p-values were significant, even without said correction. Based on the available information, there was no reason to suspect at the beginning of the semester that students participating in formative assessment at different levels would have markedly different outcomes in the course.

Table 2
Summary of Analysis of mean grade predictive variable grouped by participation level

Grade-Predictive Variable	***p* value**
ACT Math Score	.192
Cumulative GPA	.294
CRE Score	.563
Months Between Courses	.741

[1] The Calculus Readiness Exam is a multiple choice pre-calculus exam all calculus students take on the second day of class

All assessments had reliabilities within acceptable levels (Gall, Gall & Borg, 2007), and the assumptions for the statistical tests were satisfied. The limit, derivative, and definite integral labs had KR-20 values of 0.83, 0.72, and 0.78 respectively; the final exam had a Cronbach Alpha of .68. For the quantitative analyses of the whole class data, I conducted a preliminary analysis of the data to confirm that the assumptions for the statistical tests were met. The results of the normality tests for each sample used in an ANOVA appear in Table 3. Although one set of scores was not normal and these samples are not random, ANOVA is robust to these assumption violations and is still an appropriate analysis. Normality was checked using the Shapiro-Wilks Test, and all analyses were conducted with SPSS.

Table 3
Normality tests for ANOVA

	Limits, Items Discussed in Class	**Limits, Items Not Discussed in class**	**Limits, all items**	**Definite Integrals; Items Discussed in Class**	**Definite Integrals; Items Not Discussed in Class**	**Definite Integrals, All Items**
Regular	0.53	0.69	0.76	0.03	0.23	0.12
Sporadic	0.13	0.38	0.28	0.19	0.49	0.30
Nonparticipant	0.17	0.07	0.12	0.72	0.41	0.51

The data for the derivatives lab sufficiently satisfied the assumptions of an ANCOVA. For the ANCOVA analysis of the derivatives lab, there was no need to test for multicollinearity since only one covariate was used. The homogeneity of variance assumption was satisfied $(p = 0.21)$, as was the homogeneity of regression slopes $(p = 0.302)$. All of the covariate and dependent variable samples were sufficiently normal using the Shapiro-Wilks test (Table 4).

Table 4
Normality tests for ANCOVA

	Initial Submission	**Revised Submission**
Regular	0.33	0.52
Sporadic	0.45	0.56
Nonparticipant	0.12	0.75

There were 66 students that consented to participate in the study; 13 of the students were removed from the sample because they had prior exposure to the labs that could confound the results. Of the 53 students that were new to the approximation framework labs, only seven had no prior exposure to limit concepts in a prior course, and 27 of the students had AP Calculus in high school. There were 14 students classified as sporadic participants in formative assessment and 16 students classified as non-participants; the remaining 23 students participated regularly in the formative assessments (Table 5). Although students that earned A's in the course tended to be regular participants and students that failed the course tended to be nonparticipants, there was a participant at almost every combination of final grade/participation level[2], and students that earned C's in the course show no clear participation pattern.

Table 5
Final Grade by Number of Formative Assessments Completed

	A	B	C	D/F	Total
8-13	10	6	5	1	23
1-7	2	2	5	7	16
0	0	0	5	10	15
Total	12	8	16	18	54

The limits and definite integral labs were analyzed using one way ANOVAs, as was the final exam. During the derivative lab, students asked few questions, and since the previous limits labs were generally good, the discussions following each lab day were very short. However, only 8/54 students answered 14 or more items correctly on the lab write-up. Rather than recording those grades, the instructors gave all of the students that turned in a derivative lab individual written feedback on all of the questions they either answered incorrectly or left blank. The instructors told students that their first attempt would be considered a draft. Students were then given a week to revise and resubmit their derivative lab based upon the formative feedback; this became the final version of the derivative lab. In order to account for students' initial scores on the derivatives lab, an ANCOVA was used. The statistical results are given for each lab after providing a brief description of each lab and which portions were discussed in class based upon the formative post-labs.

Results

The limit lab asked students to approximate the location of a removable discontinuity where there were no obvious algebraic manipulations that would allow the discontinuity to be calculated exactly. Much of the lab depended on familiarity with function concepts. Given that there was only one formative assessment based discussion and there were no significant differences between the participation levels in any prior knowledge measure available, it is not surprising that the ANOVA found no significant differences in group achievement on the lab write-up (Table 6); the context of the lab was equally familiar to all students and there were not enough instructional interventions to make a difference. Eleven of the 20 components of the approximation framework were discussed in at least one of the postlab-based instruction sessions the class after the lab, which is indicated by asterisks in Table 6.

[2] Completion of the formative pre-lab and post-labs was 5% of the final course grade, so it was unlikely that a student would fail to complete any of these assignments and earn an A or a B in the course.

Table 6
Questions Discussed in Post-lab Based Instruction

	Contextual	Graphical	Algebraic	Numerical
Unknown Value		*		
Approximation	*	*		
Error	*	*	*	*
Error Bound	*	*	*	*
Desired Accuracy		*		

An ANOVA of student performance on the items discussed in class revealed that there was a significant difference in mean performance between at least two groups (Table 7).

Table 7
ANOVA of Items Discussed in Class, Limits Lab

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Between Groups	208.734	2	104.367	19.59	0.000
Within Groups	271.701	51	5.327		
Total	480.436	53			

The Tukey Post-hoc analysis (Table 8) showed that the regular participant group had a significantly higher mean than the other two groups, but that the sporadic and nonparticipant groups were not significantly different from each other. Given the low mean scores of these groups, this suggests that the students that were not in the regular participant group did not benefit greatly from the post-lab based instruction.

Table 8
Tukey Post-Hoc Analysis, Limits Lab (Critical Q Value = 3.44)

			Calculated Q Value	
Groups	Count	Mean Score	Sporadic	Nonparticipant
Regular	23	7.78	**7.60**	**7.10**
Sporadic	16	3.67		0.5
Nonparticipant	15	3.94		

There were three different contexts for the derivatives lab. In the first, students were asked to calculate the rate of change of surface area of a sphere whose radius was expanding, the second context dealt with radioactive decay, and most difficult context asked students to calculate the instantaneous rate of change of the gravitational force between the earth and an asteroid at a given distance. The postlabs on the derivatives lab had the lowest completion rate of any of the seven labs during the semester; this is likely because the derivatives gateway and the application of derivative chapter test were given during this lab. On the post-labs, students asked few questions, and since the previous limits labs were generally good, the discussions following each lab day were very short. However, only 8/54 students answered 14 or more items correctly on the lab write-up. Rather than recording those grades, the instructors gave all of the students that turned in a derivative lab individual written feedback on all of the questions they either answered incorrectly or left blank. The instructors told students that their first attempt would be considered a draft. Students were then given a week to revise and resubmit their derivative lab based upon the formative feedback; this became the final version of the derivative lab.

In order to investigate if there were differences in student performance after feedback, students from all participation levels that did not receive written feedback were eliminated from consideration. This left 21 regular participants, 10 sporadic participants, and five nonparticipants. One regular participant, the only regular participant that failed the course, was an outlier and eliminated from the sample, leaving 20 cases in the group.

The ANCOVA showed that there was a significant difference in mean performance on the revised derivatives lab write-ups after controlling for the score on the write-up where students received initial feedback (Table 9).

Table 9
Analysis of Covariance; Derivatives Lab

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Adjusted Means	252.17	2	126.08	6.38	.005
Adjusted Error	651.86	33	19.75		
Adjusted Total	904.03	35			

For the post-hoc analysis, I used simple contrasts with a Bonferroni correction to account for the multiple comparisons. The regular participants' mean performance was significantly higher than the mean of the non-participants. The sporadic participants' mean performance was also higher than the mean of the non-participants. Although the difference between the regular and sporadic participants is not significant, the relatively low *p*-value suggests that further exploration may be warranted (Table 10).

Table 1
Post-hoc Analysis of ANCOVA

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference Lower Bound	Upper Bound
Regular	Sporadic	2.642	1.790	0.0501	-1.005	6.288
	Nonparticipant	7.555	2.138	0.0003	3.200	11.909
Sporadic	Regular	-2.642	1.790	0.0501	-6.288	1.005
	Nonparticipant	4.913	2.292	0.0130	.245	9.581

Since students received written feedback for every incorrect or blank response on their derivatives draft, the nonparticipants received the most instructor feedback. However, even when students' initial derivatives lab write-up scores are accounted for in the ANCOVA, students in the other participation groups were significantly outperforming the Nonparticipant group on the derivative lab rewrite. This suggests that even with more extensive written feedback, the Nonparticipants were not able to increase their mean scores as much as the Regular and Sporadic participant groups did.

The definite integration labs asked students to model a given quantity with a definite integral and then approximate their quantity with Reimann sums. The contexts for the problem were volume of a portion of a sphere, mass of an object with non-constant density, probability with a continuous density function, force to stretch a spring, and water pressure on a dam. There were two formative assessment-based discussions during this lab; these discussions focused on summation notation and assistance with the technology required to calculate large Riemann sums. The ANOVA results shown in Table 11 revealed a significant difference in achievement between the three participation levels.

Table 11
Items Discussed in Post-Lab Based Instruction, Integrals Lab

	Contextual	Graphical	Algebraic	Numerical
Unknown Value			*	
Approximation			*	
Error			*	
Error Bound			*	
Desired Accuracy			*	*

The ANOVA of student performance on these six items that were discussed during the post-lab based instruction revealed significant differences in mean performance between at least one pair of groups (Table 12).

Table 12
ANOVA of Items Discussed in Post-lab Based Instruction, Integrals Lab

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Between Groups	99.531	2	49.765	63.387	0.000
Within Groups	40.04	51	0.785		
Total	139.571	53			

The Tukey Post-Hoc Analysis revealed that all three groups had significantly distinct mean performances on these items (Table 13). The three participation groups all had significantly different mean total scores on the definite integral lab: R > S (Q = 4.21), R > N (Q = 19.65), S > N (Q = 7.48). The three participation groups also all had significantly different mean scores on the items not discussed in class on the definite integral lab: R > S (Q = 9.37), R > N (Q = 19.02), S > N (Q = 9.32).

Table 13
Tukey Post-Hoc Analysis, Integrals Lab (Critical Q Value = 3.44)

			Calculated Q Value	
Groups	Count	Mean Score	Sporadic	Nonparticipant
Regular	23	4.1	**3.60**	**15.94**
Sporadic	16	3.5		**9.34**
Nonparticipant	15	1.2		

The final exam was written by the calculus instructor and course coordinator; it was administered to all introductory calculus students during a common final exam time. The cumulative common final exam ANOVA had similar results to the definite integral lab; all three groups had significantly different levels of achievement from each other, and were in the same order (Table 14).

Table 14
Results of the final exam ANOVA

	SS	**Df**	**MS**	**F**	**p**
Between	19881.53	2	9940.77	20.968	0.000
Within	24179.40	51	474.09		
Total	44059.92	53			

The post-hoc analysis (Table 15) revealed that students that never participated in the formative prelabs and post labs did significantly worse on the final than the students at the other two participation levels. The mean final exam score between students regularly participating in the formative prelabs and postlabs was not significant, but the Q value is large enough to suggest the difference in mean score between students at the regular and sporadic participation levels suggests the difference in mean score on the final exam is approaching significance.

Table 15
Tukey Post-Hoc Analysis, Integrals Lab (Critical Q Value = 3.44)

			Calculated Q Value	
Groups	Count	Mean Score	Sporadic	Nonparticipant
Regular	23	79.56	3.37	**7.39**
Sporadic	16	67.98		**4.02**
Nonparticipant	15	40.96		

The results of the study indicate that the students not participating in formative assessments are able to answer fewer questions on average than those students that do participate in the formative assessments. This is surprising because the students that did not complete any formative assessments attended class for the post-lab based instruction on the day after lab, and the students that did no formative assessments did not have significantly lower levels of prerequisites knowledge than those students participating in the formative prelabs and postlabs. The evidence is more equivocal on if the amount of formative assessments completed results in significantly higher achievement; further research is needed in this area.

Discussion

While these results indicated that there were measurable achievement differences between the three participation groups, the more interesting, and more difficult, question is why these differences exist. There appear to be two plausible explanations based on the available data. The first is that the formative assessment-based instruction was more effective for the students participating in the formative assessments. The theoretical learning trajectory for introductory calculus is that students have a great deal of trouble mastering derivatives because they tend to not have strong models for limits, but as students' model for limits improves, their achievement tends to improve (Gravemeijer & Doorman, 1999). Although the students sporadically participating in the formative assessments appear to follow this trajectory in the course, the other two participation levels do not (Figure 3); the regular participants show almost no drop in achievement between limits and derivatives.

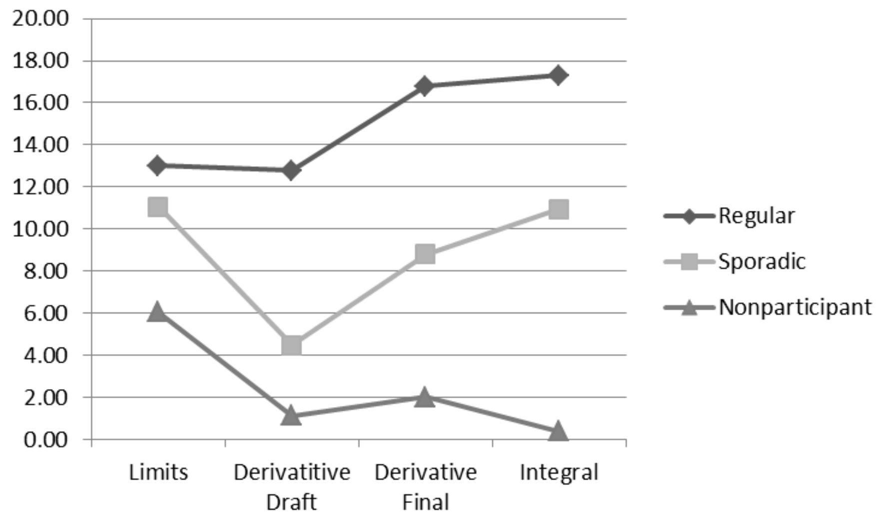


Figure 3. Achievement trajectories on the approximation labs by participation level

Another plausible explanation is these achievement patterns are indicative of a lurking variable, such as calibration differences between students at different participation levels. Calibration is considered to be a general metacognitive skill; it is the ability of a learner to accurately assess what they do and do not know (Hacker, Dunlosky, & Graesser, 1998). In this study, the opportunity for calibration occurred on the limits, first derivative lab, and the definite lab. In all three cases there was a set of questions that no student asked about on their postlabs. Since none of the students asked for help on the post-lab for these items, I considered an item to be well-calibrated if the student produced the correct solution.

In the labs, the statistical evidence for differences in calibration is not clear. In all cases the p value was smaller than 0.001, and the post hoc test showed the same significant mean differences as were found between the groups on the items discussed in class. However, there is no way to determine from the numerical data if these differences were due to differences in calibration across the groups.

Although the whole class data has no definitive answers, the qualitative case study data suggests that the three participation levels followed a similar calibration trajectory throughout the semester (Figure 2). It is likely that familiarity with the approximation framework labs accounts for much of the improvement with calibration. However, this is one of the only indications in the data I conducted that provided any insight for why the regular participants, who were not significantly better than the other students on the grade-predictive measures at the beginning of the semester, had much higher grades by the end of the semester: the regular participants maintained high calibration levels throughout the semester. Whether this was because completing formative assessments on a regular basis helped the regular participants maintain a high calibration level or if the formative assessments helped students improve their calibration throughout the semester is an area for future research.

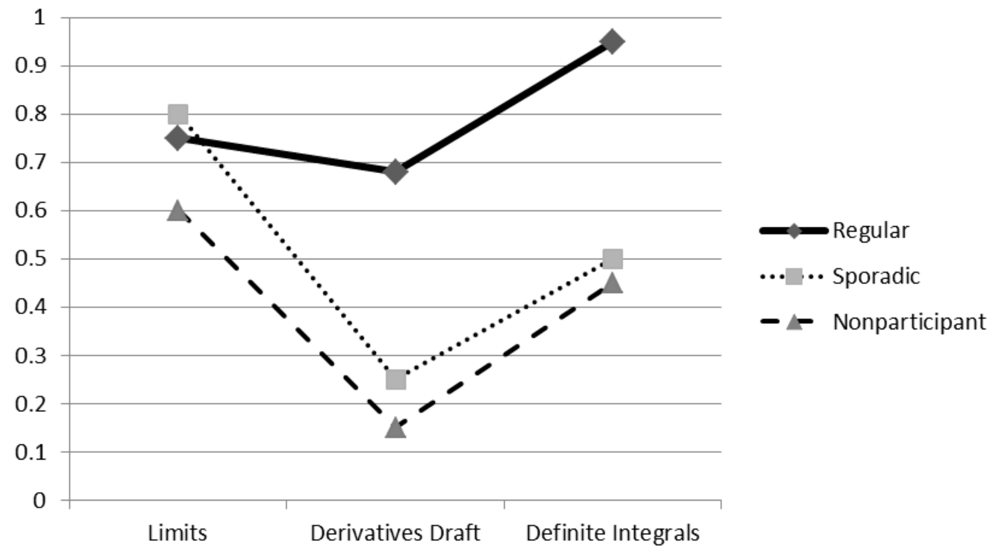


Figure 3. Percentage of well calibrated items on case study labs

Although these results are interesting, there is no data on students' initial calibration levels. There is a missing category in this analysis, since there were not enough students in this sample that completed between five and eight formative assessments to be analyzed as a separate group. More work is needed to investigate how amount of formative assessments completed affects students' achievement, and if formative assessments support students to develop better calibration skills.

References

Black, P., & McCormick, R. (2010). Reflections and new directions. *Assessment & Evaluation in Higher Education, 35*(5), 493-499.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, October 1998*, 139-148.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*(1), 5-31.

Bressoud, D. M., Carlson, M. P., Mesa, V., & Rasmussen, C. (2013). The calculus student: insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology*, (ahead-of-print), 1-15.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (in press). Meta-Analytic Methodology and Conclusions about the Efficacy of Formative Assessment. To appear in *Educational Measurement: Issues and Practices*.

Chiesa, M., & Robertson, A. (2000). Precision teaching and fluency training: making maths easier for students and teachers. *Education Psychology in Practice, 16(3)*, 297-310.

Clark, I. (2010). Formative assessment: There is nothing so practical as a good theory. *Australian Journal of Education, 54(3),* 341-352.

Clark, I. (2011). Formative assessment: Policy, perspectives and practice. Florida Journal of Education Administration and Policy, 4(2), 158-180.

DiPrete, T. A. (2013). *The Rise of Women: The Growing Gender Gap in Education and what it Means for American Schools*. Russell Sage Foundation.

Elliot, D., & Gillen, A. (2013). Images and stories: Through the eyes of at-risk college learners. *International Journal of Qualitative Studies in Education, 26(7)*, 912-931.

Gallagher, E. (2006). Improving a mathematical key skill using precision teaching. *Irish Educational Studies*, *25*(3), 303-319.

Gallagher, E., Bones, R., & Lombe, J. (2006). Precision teaching and education: is fluency the missing link between success and failure? *Irish Educational Studies, 25(1)*, 93-113.

Gall, M.D., Gall. J.P., & Borg, W.R. (2007). *Educational research* (8th ed.). New York: Pearson.

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational studies in mathematics*, *39*(1-3), 111-129.

Hacker, D. J., Dunlosky, J. E., & Graesser, A. C. (1998). *Metacognition in educational theory and practice*. Lawrence Erlbaum Associates Publishers.

Karpinski, A. C., & D'Agostino, J. V. (2012). The role of formative assessment in student achievement. *International Guide to Student Achievement*, 202.

Oehrtman, M. (2008). Layers of Abstraction: Theory and Design for the Instruction of Limit Concepts. In M. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Practice in Undergraduate Mathematics* (pp. 65-79). Washington, DC: MAA.

Oehrtman, M. (2009). Collapsing dimensions, physical limitation, and other student metaphors for limit concepts. *Journal for Research in Mathematics Education, 40,* 396–426.

Vygotsky, L. (1987). The development of scientific concepts in childhood. In R.W. Rieber and A.S. Carton (Eds.) *The collected works of L.S. Vygotsky, Volume 1: Problems of general psychology* (pp. 167-241). New York, NY: Plenium Press.

Wyatt, M. (2011). Overcoming low self-efficacy beliefs in teaching English to young learners. *International Journal of Qualitative Studies in Education, 26(2),* 238-255 DOI: 10.1080/09518398.2011.605082.

MATHEMATICS CLASSROOM OBSERVATION PROTOCOL FOR PRACTICES RESULTS IN UNDERGRADUATE MATHEMATICS CLASSROOMS

Jim Gleason
The University of Alabama

Leah D. Cofer
The University of Alabama

The purpose of this study is to determine the reliability and validity of the Mathematics Classroom Observation Protocol for Practices (MCOP^2) in undergraduate mathematics classrooms, an observation instrument designed to measure the degree to which a mathematics classroom aligns with the standards put forth by national mathematics organizations. To examine the reliability and validity of the MCOP^2 in the undergraduate setting, over thirty undergraduate mathematics classrooms at a large southeastern university were observed during the fall semester of 2013. The exploratory factor analysis conducted from the data collected indicates there are two main factors to consider in an undergraduate mathematics classroom: "lesson content" and "student engagement and classroom discourse". The internal reliability of each of these factors was verified using classical test theory to measure well at the group level.

Key words: Classroom Teaching, Evaluation, Standards

The Mathematics Classroom Observation Protocol for Practices (MCOP^2) is a K-16 mathematics classroom instrument designed to measure the degree of alignment of the mathematics classroom with the Standards for Mathematical Practice from the Common Core State Standards in Mathematics (NGACBP & CCSSO, 2010); "Crossroads" and "Beyond Crossroads" from the American Mathematical Association of Two-Year Colleges (AMATYC 1995; AMATYC 2006); the Committee on the Undergraduate Program in Mathematics Curriculum Guide from the Mathematical Association of America (Barker et al., 2004); and the Process Standards of the National Council of Teachers of Mathematics (NCTM, 2000). The instrument contains 17 items intended to measure three primary constructs (student engagement, lesson content, and classroom discourse) as validated by a review of over 150 individuals self-identified as mathematics teacher educators from a mixture of mathematics departments and departments or colleges of education (Gleason, Zelkowski, Livers, Dantzler, & Khalilian, 2014). Each of the 17 items also contains a full description of the item with specific requirements for each rating level; see Appendix B for sample items descriptions.

Purpose and Proposed Uses

Using peer reviews to evaluate faculty members' teaching effectiveness is a policy that is currently "gaining momentum" in higher education (Harris, Farrell, Bell, Devlin, & James, 2008). Seldin (1999) best defends this type of evaluation in higher education by claiming a teacher's performance in the classroom should be considered comparable to their publications and thus held to the same review process. If a faculty member's peers are supposed to reliably evaluate his or her teaching performance through classroom observations, peer review requires the "essential ingredient" of a rating scale "with scale items (that) typically address the instructor's content knowledge, delivery, teaching methods, learning activities, and the like" (Berk, Naumann, & Appling, 2004). If observing faculty members use such a reliable classroom observational protocol, peer review has potential to better measure an instructor's teaching abilities than student evaluations since there are features of a lesson that peers are better qualified to evaluate than students (Harris et al., 2008; Berk, Naumann, & Appling, 2004).

There are many universities already utilizing some version of a classroom observation protocol during their faculty peer reviews. However, these preexisting protocols are very generic, lengthy, and subjective. For instance, University of New Mexico, Tallahassee Community College, and California State University, East Bay all have observation forms online that have been adapted from *A Guide for Evaluating Teaching for Promotion and Tenure* by Centra, Froh, Gray,& Lambert (1976) where the observing faculty member can fill out an extensive forty-five item form with only three potentially biased scoring options available: "not observed", "more emphasis", or "accomplished very well" (University of New Mexico, 2006; Tallahassee Community College, 2012; California State University, East Bay, 2013; Centra et al., 1976). Since these preexisting college level protocols are not subject-matter specific, they do not necessarily draw an observer's attention toward more specific aspects of the lesson, classroom, or students "thereby resulting in potentially different kinds of teacher evaluation practice" (Spillane, Halverson, & Diamond, 2001). Spillane, Halverson, & Diamond (2001) compare 'Protocol A' consisting of a checklist of generic teaching processes with a content-specific 'Protocol B' which includes items such as "how students were required to justify their mathematical ideas" to justify why a subject-matter specific instrument would allow faculty members to identify more precise details of a teacher's performance. Instead of a generic form, a reliable content-specific observational protocol like the MCOP^2 should be used during peer observations to measure a teacher's effectiveness and thus help generate a discussion on quality teaching in college undergraduate classrooms throughout the United States.

This review process is also useful for generating discussion among future and current college mathematics faculty about teaching. It can be used to help new graduate students better prepare for the classroom, help departments to decide goals for teaching and how well those goals are being met, and be used while observing classes as a group to generate discussion about what the standards look like in the college setting.

Since the MCOP^2 is grounded in the national recommendations of organizations focused on post-secondary education, it is useful to explore the current practice of mathematics teaching and its relationship to student learning. This instrument allows for a quantification of different aspects of college mathematics teaching that could then be used to explore teachers' choices in the classroom, effects of different teaching styles with different types of students, how a teacher's practice in the classroom changes with the different topics and situations throughout a semester, and many more.

Preexisting Protocols

There are many content-specific classroom observation protocols already available for use in elementary, middle, and high school mathematics classes (Hill, Charalambous, Blazar, et al., 2012). Three existing classroom protocols claim to extend to college level mathematics classrooms, and while each of these protocols is described as unique, all three credit Horizon Research Corporation, Inc. for their development (Weiss, Pasley, Smith, Banilower, & Heck, 2003; Wainwright, Morrell, Flick, & Schepige, 2004; Walkington et al., 2012; Sawada et al., 2000a). Unlike the MCOP^2, these preexisting protocols are not designed specifically for mathematics classrooms, but instead are intended for use in both mathematics and science classrooms (Wainwright, Flick, & Morrell, 2003; Walkington et al., 2012; Sawada et al., 2000a). In order to maintain this dual purpose, logically these protocols use science terminology within some of their protocol descriptors such as "Students made predictions, estimations, and/or hypotheses and devised means for testing them" (Sawada et al, 2000b), making it difficult for an observer of a mathematics classroom to definitively score certain items. While some of the preexisting protocols have claimed to test for predictive validity in college mathematics classrooms (Sawada et al., 2000b), no preexisting protocol has done a study including more than a few strictly undergraduate mathematics classrooms so no preexisting protocol has actually proven its reliability or validity in college mathematics classrooms.

One of the most widely used classroom observational protocols in public school mathematics classes is the Reformed Teaching Observation Protocol (RTOP), developed by the Evaluation Facilitation Group of the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT) (Sawada et al., 2000a). While the RTOP is widely praised for its reliability and validity in both math and science public school classrooms, surprisingly the RTOP Reference Manual contains very few references to articles in mathematics education yet numerous references to articles in science education. Out of the seventeen references listed, six references are strictly for science education, four articles are on learning and the brain, two articles are the RTOP's first Technical Report and Training Guide, one article is from the 1980s on both math and science education, and the remaining four are citations to various years of NCTM standards from 1989 – 2000 (Sawada et al., 2000b). The creators of the RTOP admit in the "Test Development" section of their Reference Manual that the language of the items in the first draft of the instrument was "particularly referenced toward science teaching" and thus hard to interpret in a mathematics classroom (Sawada et al., 2000b). Mathematicians in the ACEPT project critiqued and suggested making "an unequivocal request to overhaul the science-dominated language"; therefore, a mathematics educator then modified the wording of items on the instrument without changing the original structure (Sawada et al., 2000b). While the RTOP in commonly used in both math and science public school classrooms, its references and item language make it better geared for use in science classes.

In addition, the instrument was originally tested in 13 "introductory" mathematics classes at universities and community colleges, but that is a substantially small portion of the 153 total classrooms that participated, particularly when compared to the 63 college level science classes observed (see Table 12, Sawada et al., 2000b). Furthermore, it is interesting to note that the college classroom teacher samples consisted of "a large number of faculty who were involved in the ACEPT initiative", thus the authors concluded this could be a reason why the college classrooms samples had higher scores on the RTOP than the middle and high school samples (Sawada et al., 2000b). In 2002, the ACEPT program tried to extend their method of reformed teaching to the college level by attempting to "incorporate reformed teaching methods in several nonmajors' and majors' courses"; hence, the RTOP was again tested in certain college classrooms. However, the only mathematics course observed was "Theory of Elementary Mathematics", a course designed specifically for preservice elementary school teachers (Lawson et al., 2002). Unlike the MCOP^2, the RTOP does not utilize the most recent national standards for mathematics classrooms and the RTOP has not extensively been tested in strictly mathematics college level classes taught by ordinary faculty members.

A classroom observation protocol that supposedly extends to college level mathematics classrooms is the Oregon-Teacher Observation Protocol (O-TOP), created by the Oregon Collaborative for Excellence in the Preparation of Teachers (OCEPT) as part of the Collaborative for Excellence in Teacher Preparation program of the National Science Foundation (Wainwright et al., 2003). According to Wainwright et al. (2004), "A major focus of the OCEPT grant was to engage science and mathematics faculty members teaching undergraduate courses in institutions across the state in a critical examination of their instructional practices." Even though the O-TOP was allegedly designed for use in both science and mathematics classrooms, their only citations strictly pertaining to mathematics are to various years of NCTM standards from 1989 – 2000 (Wainwright et al., 2004). Furthermore, the only mathematics classrooms observed were courses taught by OCEPT Faculty Fellows, so these were not typical college mathematics classes: "Of the 10 mathematics observations, two were lecture, one was lecture with discussion, and the remaining seven were small group discussion" (Wainwright et al., 2004). Despite its supposed reliability in Faculty Fellows mathematics classes, the OTOP's scientific nature and lack of recent mathematical standards make it undesirable for use in college mathematics courses.

Another classroom observation protocol supposedly appropriate for use in mathematics and science classrooms "from kindergarten to college" is the UTeach Observation Protocol (UTOP) created by the UTeach program at the University of Texas at Austin (Walkington et al., 2012). The protocol was developed to evaluate UTeach graduates, particularly Noyce Scholars, in order to fulfill a National Science Foundation requirement (Walkington & Marder, 2013). The language used within the UTOP demonstrates it is extremely science-based. For instance, an indicator in the protocol on the pace and flow of the lesson has a science-specific example listed with it in the Training Guide: "e.g. most of a science lab is focused on directions instead of content development" (Marder et al., 2010). Besides the science-specific language, another drawback to the UTOP is it is solely based off of NCTM standards from 1991 (Walkington et al., 2012). Furthermore, even though its authors originally planned to conduct observations at the college level in order to refine the instrument, no study has documented testing the UTOP in an undergraduate mathematics classroom (Walkington et al., 2012). The UTOP's lack of testing at the college level, along with its standards deficiency and overuse of scientific language, show its inapplicability to college mathematics classrooms.

MCOP^2 Framework

Since the MCOP^2 was formed for mathematics-specific classrooms using Common Core State Standards in Mathematics (CCSSM) *Standards for Mathematical Practice*, the American Mathematical Association of Two-Year Colleges' *Crossroads* and *Beyond Crossroads,* the Mathematical Association of America's *CUPM Curriculum Guide*, and the latest NCTM standards, it is the only classroom observation protocol available that is applicable for use in K-16 mathematics instruction. After the nation adopted the CCSSM for use in public school mathematics classrooms, the Association of Public and Land-grant Universities (APLU) issued a brief laying out an "action agenda" with four main points for the role of higher education institutions (APLU, 2011). Point one addresses the issue of aligning curriculum between K-12 and higher education, and the APLU later indicates disciplinary departments should be "transforming introductory courses so that they are aligned with CCSSM (in both content and approach)" (APLU, 2011; King, 2011). College teachers themselves agree that their curriculum should be aligned with CCSSM since during a study on the applicability of the Common Core State Standards, over 1800 instructors found the Standards for Mathematical Practice to be extremely applicable and important to their courses (Conley, Drummond, de Gonzalez, Rooseboom, & Stout, 2011). Furthermore, point three of the APLU's agenda states higher education institutions should be "conducting research on issues of teaching and learning the Common Core State Standards, teacher quality, and the implementation of the Common Core State Standards" (APLU, 2011; King, 2011). Thus there is a need for a CCSSM-based mathematics classroom protocol and the MCOP^2 is the only protocol intentionally designed to meet this requirement.

Each of the items on the MCOP^2 was designed to coordinate with a Standard for Mathematical Practice, and in turn thus correlates to a recommendation in the CUPM Curriculum Guide. For instance, Item #9 on the protocol is "The lesson provided opportunities to examine elements of abstraction (symbolic notation, patterns, generalizations, conjectures, etc.)," matching the second Standard for Mathematical Practice that instructors should be aiming to teach their students**: "**CCSS.Math.Practice.MP2: Reason abstractly and quantitatively" (NGACBP & CCSSO, 2010). This concept also connects to Part 1 of the CUPM Curriculum Guide which gives recommendations for departments, programs, and all courses by Barker, et al. (2004): "For instance, one reason students encounter difficulty in applying mathematics to problems in other disciplines is that they have trouble identify appropriate mathematical procedures when problems are expressed with different symbols than those used in the mathematics classroom….instructors can go beyond conventional x, y notation to use a larger collection of symbols for both constants and variables." (p. 20)

Therefore, both the CCSSM and CUPM specifically address this important aspect of a teacher's lesson content which Item #9 is designed to measure. This correlation between the teacher and student behaviors detected by the MCOP^2, the Standards for Mathematical Practice, and the CUPM Curriculum Guide extends to all seventeen items on the protocol.

Methodology

A pilot study to field test the MCOP^2 in undergraduate mathematics classrooms was implemented during the fall semester of 2013, and observations by the research team composed of a graduate student in mathematics and a mathematics professor were scheduled based upon instructor approval. Twenty-eight of the fifty-eight teachers agreed to participate in this initial study. Since some of these faculty members teach two completely different courses at the university, a total of thirty-six classroom observations occurred throughout the semester.

From the 36 classrooms participating in the study, 15 classes were taught by Graduate Teaching Assistants, 8 classes were taught by non-tenure faculty, and 13 classes were taught by tenured or tenure-track. There was a diverse amount of courses in this sample, ranging from college algebra to upper division mathematics. In the norm section of the results, the observed classes are grouped into five main categories: Precalculus, which includes college algebra to algebra with trigonometry courses; Applied Calculus which is a business calculus course; Calculus including Calculus I and II, differential equations, and computationally focused introductory linear algebra courses; Education which contains mathematics courses specifically designed for preservice elementary and secondary mathematics teachers; and Proof consisting of upper division proof-based mathematics courses. This study's observations included 11 Precalculus classes, 5 Applied Calculus classes, 11 Calculus classes, 4 Education classes, and 5 Proof classes. To determine the structure and reliability of the instrument, each class was observed once during the semester, and the analysis of the data collected from these thirty-six completed MCOP^2 forms is analyzed using exploratory factor analysis and classical text theory analysis.

Results

The seventeen item MCOP^2 was analyzed using observations from thirty-six undergraduate mathematics classrooms. The researchers originally anticipated three factors to appear in the analysis with each factor corresponding to one of the three sections of the instrument (Student Engagement contains Items 1-5, Lesson Content contains Items 6-11, Classroom Culture and Discourse contains Items 12-17) (Gleason et al., 2014). However, as shown by the Scree Plot below (Figure 1), there are actually only one or two applicable factors. The third potential factor is not a legitimate component, despite its arguable location in the curvilinear region, since it is such a low eigenvalue. Furthermore, the Factor Matrix of a potential 3-Factor Model indicated the items from both Student Engagement and Classroom Culture and Discourse were actually loading onto the same factor, henceforth called Student Engagement and Classroom Discourse.

Figure 1: Scree Plot of Entire Protocol

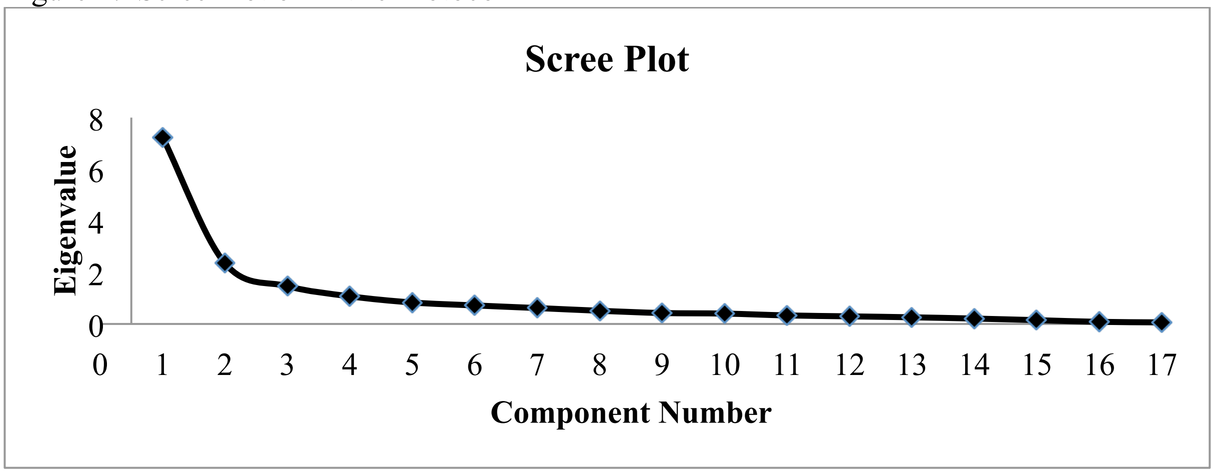
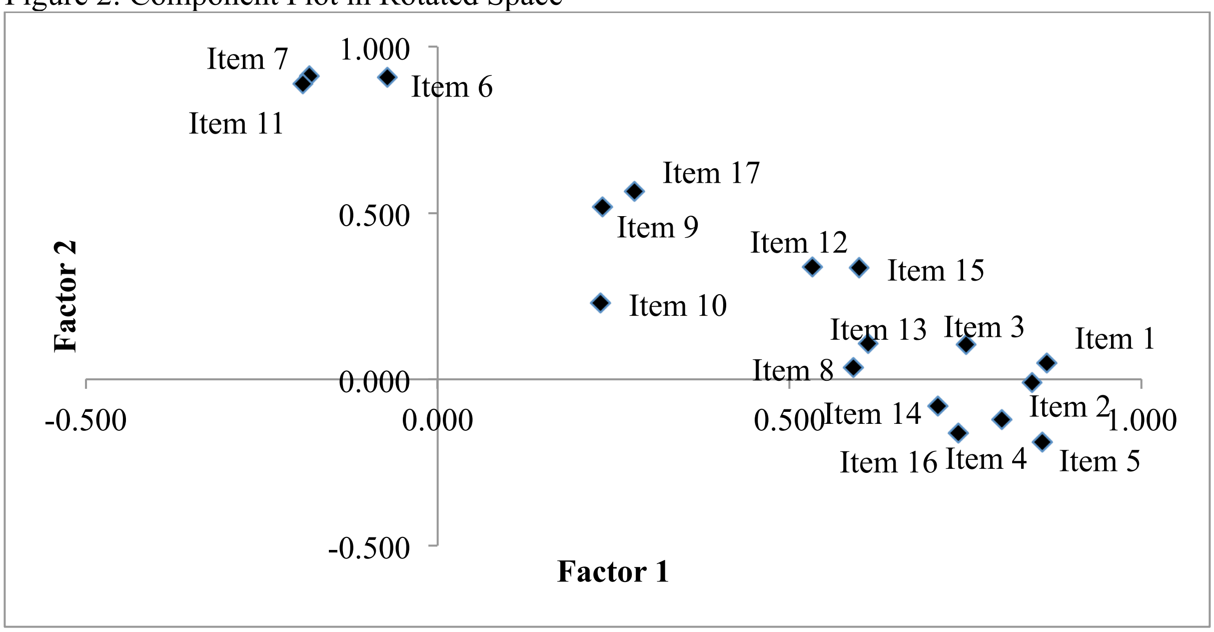


Figure 2: Component Plot in Rotated Space

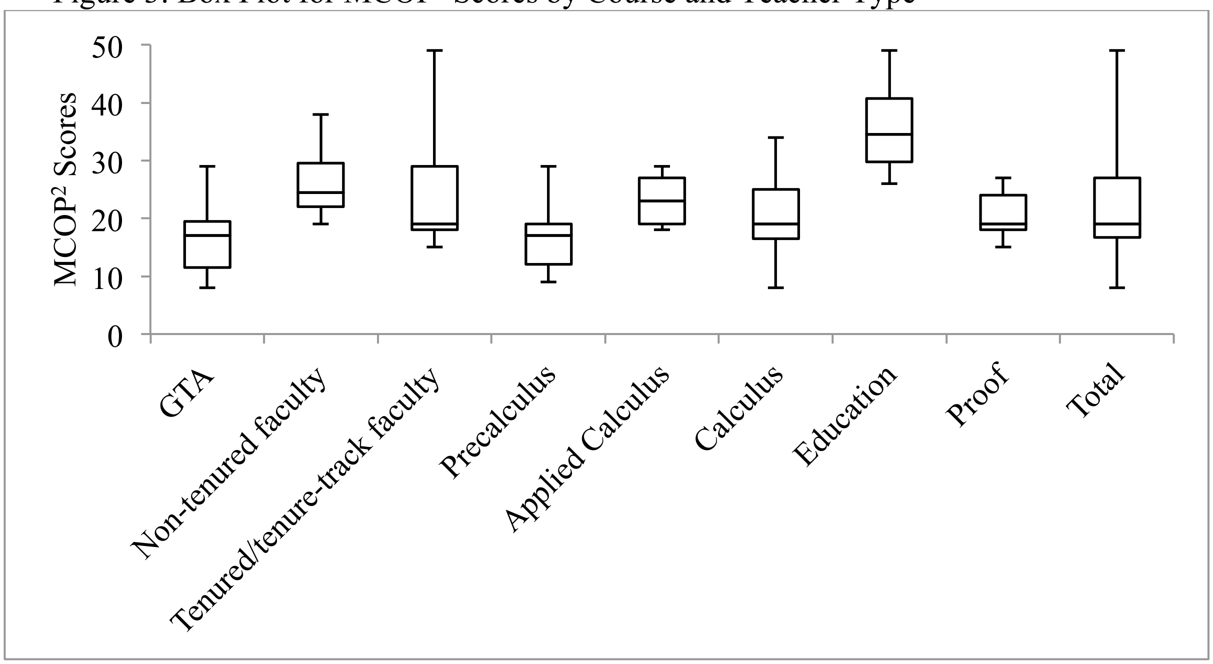


Solutions asking for two principle components to be extracted resulted in a 2-Factor Model explaining over 50% of the total variance. Using a promax rotation with Kaiser normalization, the component plot (Figure 2) indicates Items 1-5 and 12-16 are correlated, as expected, to Factor 1 (Student Engagement and Classroom Discourse) and Items 6, 7, 9, 10, and 11 are correlated to Factor 2 (Lesson Content). Items 8 and 17 did not load as expected, but instead loaded on the opposite factor. Since Item 8 was “The lesson promoted modeling in mathematics”, it does not fit the theoretical construct of Student Engagement and Classroom Discourse, and so was removed from the reliability analysis of that subscale, but included on the overall instrument. Since Item 17 (The teacher uses student questions/comments to enhance mathematical understanding) fits well within the construct of lesson content, the results from this item are included in that subscale reliability, as well as the entire instrument.

The subscales of “Lesson Content” and “Student Engagement and Classroom Discourse” were both found to be unidimensional with over 50% of the variance contained in a single factor and so can be treated as subscales and analyzed for their own Cronbach’s alpha reliabilities of 0.779 and 0.907, respectively. In addition, if one considers the entire instrument to be unidimensional, then the entire protocol has a Cronbach’s alpha of 0.898. Therefore, the internal reliabilities are high enough for both subscales and the entire instrument to be used to measure at the group level, either multiple observations of a single classroom or single observations of multiple classrooms.

Norms from the 36 classroom sample used to create the factor analysis above are shown in Figure 3 to give future users of the MCOP^2 some standards of performance against which to assess the scores achieved by individuals or samples in their own data sets. In addition, from these norms, one can see that the instrument was able to differentiate between types of instructors and types of classes.

Figure 3: Box Plot for MCOP^2 Scores by Course and Teacher Type



Conclusions

The overall instrument’s high coefficient alpha of .898 is noteworthy in that it demonstrates that the instrument is measuring something and is able to differentiate between classroom settings. Furthermore, the exploratory factor analysis indicates the MCOP^{2}gauges two main factors in an undergraduate mathematics classroom: “Lesson Content” and “Student Engagement and Classroom Discourse”. Thus when the instrument is separated into two sections, the MCOP^2’s Student Engagement portion demonstrates an exceptionally high level of internal reliability, proving the instrument successfully gauges an undergraduate mathematics classroom’s culture and student participation. Although not as remarkably high as the Student Engagement portion, the MCOP^2’s Lesson Content portion also shows high internal reliability, indicating the instrument also successfully measures the content of a mathematics lesson in a college level classroom.

This initial study has provided results indicating the Mathematics Classroom Observation Protocol for Practices (MCOP^2) is a reliable observational protocol for undergraduate mathematics classrooms. However, a much larger study of the instrument's reliability at the college level by testing the instrument in undergraduate mathematics classrooms at multiple higher education institutions, is needed as the data collected from observations at numerous community colleges, liberal arts schools, and other research universities would better examine and solidify the MCOP^2's structure and reliability in a more general college mathematics classroom setting.

References

AMATYC (1995). *Crossroads in Mathematics: Standards for Introductory College Mathematics before Calculus.* (D. Cohen, Ed.) Memphis, TN: American Mathematical Association of Two-Year Colleges. Retrieved from http://www.amatyc.org/?page=GuidelineCrossroads

AMATYC (2006). *Beyond Crossroads: Implementing Mathematics Standards in the First Two Years of College.* (R. Blair, Ed.) Memphis, TN: American Mathematical Association of Two-Year Colleges. Retrieved from http://beyonecrossroads.matyc.org/

APLU (2011). *The Common Core State Standards and Teacher Preparation: The Role of Higher Education*. APLU/SMTI, Paper 2. Washington, DC: Association of Public and Land-grant Universities. Retrieved from http://www.aplu.org/document.doc?id=3482

Barker, W., Bressoud, D., Epp, S., Ganter, S., Haver, B., & Pollatsek, H. (2004). *Undergraduate Programs and Courses in the Mathematical Sciences: CUPM Curriculum Guide, 2004*. Mathematical Association of America. 1529 Eighteenth Street NW, Washington, DC 20036-1358.

Berk, R. A., Naumann, P. L., & Appling, S. E. (2004). Beyond student ratings: Peer observation of classroom and clinical teaching. *International Journal of Nursing Education Scholarship*, *1*(1), 1-26.

California State University, East Bay. (2013). *Classroom observation worksheet.* Retrieved from http://www20.csueastbay.edu/faculty/ofd/resources/archive/class-worksheet-centra.pdf

Centra, J. A., Froh, R. C., Gray, P. J., & Lambert, L. M. (1976). A guide for evaluating teaching for promotion and tenure. Syracuse, NY: Syracuse University Center for Instructional Development.

Conley, D. T., Drummond, K. V., de Gonzalez, A., Rooseboom, J., & Stout, O. (2011). Reaching the Goal: The Applicability and Importance of the Common Core State Standards to College and Career Readiness. *Educational Policy Improvement Center*.

Gleason, J., Zelkowski, J., Livers, S., Dantzler, J., & Khalilian, J. (2014). Mathematics classroom observation protocol for practices: Validity and reliability. Preprint.

Harris, K-L., Farrell, K., Bell, M., Devlin, M., and James, R. (2008). Peer review of teaching in Australian higher education: A handbook to support institutions in developing effective policies and practices.

Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., Humez, A., Litke, E., & Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment*, *17*(2-3), 88-106.

King, J. E. (2011). Implementing the Common Core State Standards: An Action Agenda for Higher Education. *State Higher Education Executive Officers*.
Lawson, A. E., & Bloom, I., Falconer, K., Hestenes, D., Judson, E., Piburn, M.D., Sawada, D., Turley, J., Wyckoff, S. (2002). Evaluating college science and mathematics instruction. *Journal of College Science Teaching*, *31*(6), 388-393.
Marder, M., Abraham, L., Allen, K., Arora, P., Daniels, M., Dickinson, G., Ekberg, D., Walker, M. (2010). *The UTeach Observation Protocol (UTOP) training guide (adapted for video observation ratings)*. Austin, TX: UTeach Natural Sciences, University of Texas Austin.
NCTM. (2000). *Principles and Standards for School Mathematics*. Washington, D.C.: National Council of Teachers of Mathematics.
NGACBP & CCSSO. (2010). *Common Core State Standards for Mathematics.* Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers.
Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000a). Reformed Teaching Observation Protocol (RTOP) (ACEPT Technical Report No. IN00-1). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.
Sawada, D., Piburn, M., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000b). Reformed teaching observation protocol (RTOP) reference manual. *Tempe, Arizona: Arizona Collaborative for Excellence in the Preparation of Teachers*.
Seldin, P. (1999). Current practices – good and bad – nationally. In P. Seldin & Associates (Eds.), Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions (pp. 1–24). Bolton, MA: Anker.
Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational researcher*, *30*(3), 23-28.
Tallahassee Community College. (2012). *Classroom observation worksheet.* Retrieved from https://www.tcc.fl.edu/FacultyStaff/CTLL/Documents/Classroom_Observation_Worksheet.pdf
University of New Mexico. (2006). *Classroom observation report.* Retrieved from http://www.unm.edu/~oset/SupportingDocuments/example%20classroom%20observation%20forms.pdf
Wainwright, C., Flick, L., & Morrell, P. (2003). The development of instruments for assessment of instructional practices in standards-based teaching. *Journal of Mathematics and Science: Collaborative Explorations*, *6*(1), 21-46.
Wainwright, C., Morrell, P. D., Flick, L., & Schepige, A. (2004). Observation of reform teaching in undergraduate level mathematics and science courses. *School Science and Mathematics*, *104*(7), 322-335.
Walkington, C., Arora, P., Ihorn, S., Gordon, J., Walker, M., Abraham, L., & Marder, M. (2012). Development of the UTeach observation protocol: A classroom observation instrument to evaluate mathematics and science teachers from the UTeach preparation program. Retreived from https://uteach.utexas.edu/sites/default/files/UTOP_Paper_Non_Anonymous_4_3_2011.pdf
Walkington, C., & Marder, M. (2013). Classroom Observation and Value-Added Models Give Complementary Information about Quality of Mathematics Teaching. Retrieved from https://uteach.utexas.edu/sites/default/files/WalkingtonMarderMET2013.pdf

Weiss, I., Pasley, J., Smith, P. S., Banilower, E., & Heck, D. (2003). Looking inside the classroom: a study of K-12 mathematics and science education in the United States. Horizon Research Inc., Chapel Hill, NC.

Appendix A: MCOP^2 Items

1. Students engaged in exploration/investigation/problem solving.
2. Students used a variety of means (models, drawings, graphs, concrete materials, manipulatives, etc.) to represent concepts.
3. Students were engaged in mathematical activities.
4. Students critically assessed mathematical strategies.
5. Students persevered in problem solving.
6. The lesson involved fundamental concepts of the subject to promote relational/conceptual understanding.
7. The lesson promoted connections across the discipline of mathematics.
8. The lesson promoted modeling with mathematics.
9. The lesson provided opportunities to examine mathematical structure. (symbolic notation, patterns, generalizations, conjectures, etc.)
10. The lesson included tasks that have multiple paths to a solution or multiple solutions.
11. The lesson promoted precision of mathematical language.
12. The teacher's talk encouraged student thinking.
13. There were a high proportion of students talking related to mathematics.
14. There was a climate of respect for what others had to say.
15. In general, the teacher provided wait-time.
16. Students were involved in the communication of their ideas to others (peer-to-peer).
17. The teacher uses student questions/comments to enhance mathematical understanding.

Appendix B: Sample MCOP^2 Item Descriptors

5. Students persevered in problem solving.

One of the Standards for Mathematical Practice (NGACBP & CCSSO, 2010) is that students will persevere in problem solving. Student perseverance in problem solving is also addressed in the Mathematical Association of America's Committee on the Undergraduate Program in Mathematics Curriculum Guide (Barker et al., 2004): "Every course should incorporate activities that will help all students…approach problem solving with a willingness to try multiple approaches, persist in the face of difficulties, assess the correctness of solutions, explore examples, pose questions, and devise and test conjectures."

Perseverance is more than just completion or compliance for an assignment. It should involve students overcoming a road block in the problem solving process.

Score	Description
3	Students exhibited a strong amount of perseverance in problem solving. The majority of students looked for entry points and solution paths, monitored and evaluated progress, and changed course if necessary (NGA & CCSSM, 2010; Barker et al., 2004). When confronted with an obstacle (such as how to begin or what to do next), the majority of students continued to use resources (physical tools as well as mental reasoning) to continue to work on the problem.

2	Students exhibited some perseverance in problem solving. Half of students looked for entry points and solution paths, monitored and evaluated progress, and changed course if necessary (NGA & CCSSM, 2010; Barker et al., 2004). When confronted with an obstacle (such as how to begin or what to do next), half of students continued to use resources (physical tools as well as mental reasoning) to continue to work on the problem.
1	Students exhibited minimal perseverance in problem solving. At least one student but less than half of students looked for entry points and solution paths, monitored and evaluated progress, and changed course if necessary (NGA & CCSSM, 2010; Barker et al., 2004). When confronted with an obstacle (such as how to begin or what to do next), at least one student but less than half of students continued to use resources (physical tools as well as mental reasoning) to continue to work on the problem. There must be a road block to score 1 -3.
0	Students did not persevere in problem solving. This could be because there was no student problem solving in the lesson, or because when presented with a problem solving situation no students persevered. That is to say, all students either could not figure out how to get started on a problem, or when they confronted an obstacle in their strategy they stopped working.

7. The lesson promoted connections across the discipline of mathematics.

This item focuses on helping students to see connections between different parts of mathematics. For early elementary grades, this could be a connection between measurement and counting or area models for multiplication. In the middle grades, this could be a connection between area and distributive property, or a connection between operations on different number systems. At the high school level an example would be connections between algebraic and geometric reasoning, or a connection between the different types of inverses. In an undergraduate classroom, this could be an opportunity for students to explore mathematical ideas from a variety of perspectives, or a connection to other subjects (both in and out of the mathematical sciences), or a connection to a contemporary topic from the mathematical sciences and its applications.

Score	Description
3	Connections are emphasized throughout the lesson and/or are a major component of the lesson.
2	Connections are frequent throughout the lesson, but the connections are not a major component of the lesson.
1	A few connections are made in the lesson, but it is not frequent.
0	The lesson just makes no connections between mathematical topics.

A Framework for Characterizing Students' Thinking about Logical Statements and Truth Tables

Casey Hawthorne
San Diego State University

Chris Rasmussen
San Diego State University

While a significant amount of research has been devoted to exploring why university students struggle applying logic, limited work can be found on how students actually make sense of the notational and structural components used in association with logic. This project borrows the theoretical framework of unitizing and reification, which have been effectively used to explain the types of integrated understanding required to make sense of symbols involved in numerical computation and algebraic manipulation, to investigate students' conceptualizations of truth tables and implication statements. We use a continuum as a framework to analyze the degree to which students' thinking of each is compartmentalized versus unified. Results indicate that students tend to treat the constituent pieces that make up these mechanisms independently without an understanding of each as a whole or an integrated view of the two together. Consequently, students manipulate symbols without an appreciation of the overarching meaning.

Keywords: Unitizing, Reification, Logic, Truth Tables, Proof

Introduction

For many mathematicians, the essence of mathematics is captured in proofs (Ross, 1998; Rav, 1999). Despite the centrality of proof, however, its role in school mathematics in the United States is peripheral at best, with its only substantial treatment in the secondary mathematics curriculum occurring in a one-year geometry course (Moore, 1994; Wu, 1996, Knuth, 2002). To reverse this trend, recent reform efforts have significantly elevated the status of proof in school mathematics (National Council of Teachers of Mathematics [NCTM], 2000; Common Core State Standards in Mathematics, 2012). Still, for most students going through the K-12 US curriculum and even into the first two years of post-secondary studies, mathematics involves predominantly carrying out various procedures in order to solve numerically and algebraically based problems. Students might be exposed to the idea of proof when a teacher presents and explains a theorem that will be used later by the class, but the emphasis is on mastery of the result, not on the process that gives rise to the conclusion or its comprehension (Herbst, 2002). Consequently, it is no surprise that students struggle significantly when the mathematical focus changes in upper-level university mathematics courses from computation based activity to more abstract deductive reasoning and the communication of these ideas. To help bridge this gap and provide students with the requisite skills to successfully generate and make sense of mathematical arguments on their own, many universities have begun to incorporate the explicit instruction of proof into their curriculum, creating "transition to proof" courses (Moore, 1996).

A common component in transition to proof courses is the introduction of formal logic. The inclusion of formal logic is sensible for two interrelated reasons. First, in order for students to write a proof, they must be aware of what is necessary to establish whether a statement is true or false (Epp, 2003). The rules that govern formal mathematical arguments and mathematical statements are different from those used in informal everyday speech. For example, in common use the statements, "There is a mother for all children" and "All children have a mother," are commonly used interchangeably inferring the later mathematical meaning (Dubinsky & Yiparaki, 2000). Teaching logic makes clear to students the exact conventions, approaches, and methods employed by the mathematical community and it highlights the precision of mathematical thought. Second, logic provides a framework and language to communicate the exact requirements necessary for a particular mathematical argument. As such, it can be viewed as a tool that helps students break down complex statements and identify the precise features and properties that must be established in order to prove a claim. For example, mastering the rules of negation, students are able to recognize the specific properties involved in producing a counterexample. Epp (2003) highlights that logic, in particular logic tables, can act as a scaffolding tool, making the abstract nature of argumentation more tangible. Logic tables create a structure to help students "organize their knowledge about logical principles and gives them concrete objects to hang onto while they deal with the abstraction of the logic" (Epp, 2003, p. 986).

As compelling as the instruction of logic appears, there is evidence that such a focus results in limited to no improvement in student achievement. For example, Cheng, Holyoak, Nisbett, and Oliver (1986) found no difference in performance on modus ponens and modus tollens tasks between university students who had taken an introductory logic course and a control group of students who had not. Consequently, they hypothesized that students do not reason purely abstractly. As participants did show improvement after training in specific functional categories of conditional statements, they conjectured that students use domain dependent schemas instead. Another explanation for this unexpected result is that students have difficulty connecting the syntactic rules of logic to real life applications. Selden & Selden (1995) found that students often struggle decoding statements written with a more familiar, colloquial structure and translating them into formal mathematical language. Unable to link the two domains, students are unable to tap into the power of formal logic.

The broad goal of this study was to explore how students coordinate the meaning between contextualized statements and the corresponding formal logical symbolic representations. What emerged in the analysis, however, was a different phenomenon. As students were asked to operate on symbolically written logical statements and comment on their associated contextual interpretation, their explanations revealed how they thought about and made sense of the logical tools themselves, which gave rise to the following specific research question: How do students make sense of the symbolic expression $p \Longrightarrow q$ and its related truth table? In answering this question, we developed a framework for interpreting and characterizing how students make sense of logical statements.

Theoretical Background

Theorists in mathematics education have put forward various characterizations for how mathematical concepts are understood and how mathematical reasoning develops. For example, in the early 80's Steffe (1983) proposed that unitizing and the resultant formation of a composite unit constitutes a robust understanding of number. A child is said to view number, such as 12, as a composite unit if he or she can simultaneously treat 12 as 12 individual units or as a single entity composed of 12 units and be able to partition 12 in multiple ways (e.g., as 10 and 2). Lamon (1996) extended this notion into rational number and described unitizing more generally as the cognitive assignment of multiple mathematical entities into a combined whole. This newly constructed abstract object can then be used to reason with as a single unit. She emphasized that the key with unitizing is the ability to connect multiple pieces and envision them as a single, collective chunk, while at the same time retaining an appreciation of their individual parts relative to each other and to the newly conceptualized whole. Such a mathematical conceptualization leads to more sophisticated thinking as "it allows students to think about both the aggregate and the individual items that compose it" simultaneously (Lamon, 1996, p. 171).

Wheatley and Reynolds (1996) further expanded the notion of a composite unit beyond number contexts to analyze students' approach to a geometric tiling problem. In their study, students who combined geometric shapes to create and conceive of the geometric units as a single unit were far more successful. Similar to the observed cases with numbers, these students demonstrated the ability to iterate these shapes in productive ways, flexibly partition them, as well as describe various resulting patterns. Having generated their own composite unit, they demonstrated the ability to simultaneously see and manipulate the shape as a single whole as well as the individual comprising parts.

The notion of a composite unit is closely related to the construct of reification, which explains how mathematical constructs can be understood in two different ways. On the one hand, an individual may think about a mathematical idea in an operational way, where notation is viewed as a set of instructions for a particular process (Sfard & Linchevski, 1994). For example, the algebraic expression $3(x + 5) – 1$ might be interpreted as a collection of procedures, specifying that for any number the operations "add 5", multiply by 3", and then "subtract 1" be applied. On the other hand, algebraic expressions can embody a structural conception and represent the result of these processes. With this interpretation, the process has been reified and each of the various computations is considered as a whole unified and completed object. As such, this newly created mathematical entity can then be treated as a single entity, or in other words, as a composite unit.

As Sfard (1995) points out, this dual role enables mathematical notation to be an extremely powerful tool. It allows the user to understand and conceptualize a very complex and involved process, while at the same time treat it as a single entity. It can then be manipulated and simplified syntactically, without the large burden that the operational mode of thinking places on working memory. This capacity, though, can also act as a double edged sword, what Sfard and Linchevski (1994) refer to as a pseudostructural conceptualization. Often students are introduced to powerful symbolic notation along with various procedures to apply to them, but fail to develop an underlying grasp of the processes the notation embodies. As Sfard (1995) highlights, notational expressions become viewed as "meaningless symbols governed by arbitrary established transformations" (p. 30). In the end, the manipulation itself becomes the focus of the activity and the symbolic results are seen as producing the answer themselves.

Through the development and use of various symbolically based tools, the richness of logic is now able to be represented in compact, easy to manipulate representations. Unfortunately, if students compartmentalize different notational pieces of logical symbols without simultaneous reference to the composite whole, or treat the symbols without reference to the semantics they represent, they fail to appreciate the different layers and meanings which the symbols embody. In our analysis of how students make sense of the symbolic expression $p \Longrightarrow q$ and its related truth table, we bring to bear the related notions of composite unit and reification to interpret and characterize student thinking. In particular, both of these lenses point to how the expression $p \Longrightarrow q$ and its related truth table might be understood in a compartmentalized way or in a unified way.

Methods

The participants for this study were all drawn from a discrete mathematics course at a large university in the southwest of the United States. The course served two main populations and attempts to balance the achievement of two major goals. Roughly half the students were mathematics majors while the other half were computer science majors. The curriculum aimed to expose students to the mathematical content ordinarily associated with discrete mathematics such as set theory, logic, combinatorics and graph theory, while at the same time introducing and exposing them to the fundamental elements of mathematical proof and communication. Thus, this course functioned as a transition to proof course, using discrete mathematics to illustrate the precision of mathematical definitions and the rigorous methods for establishing the truth of mathematical statements. The class size had approximately 80 students enrolled (with only half attending on a regular basis), and instruction followed more a modified standard lecture format where students were encouraged to ask questions and time was given on occasion for students to work through specific problems.

Six students volunteered to participate in individual, think aloud problem solving interviews. The six students, while exhibiting different degrees of academic achievement during the class, were all highly successful, scoring above average on all assessments and demonstrating more or less mastery of topics on assessments. Zach and Alan received an A for the course, Kate and Eduardo received high B's/low A's, and Sofia and Cody received B's. They were described by the professor as engaged and hardworking students, and regular classroom visits made by the first author confirmed the instructor's view. Each attended class regularly and actively participated through diligent note taking and asking questions in class. As such, our analysis highlights the mathematical thinking of the more engaged and successful students.

Each student participated in a 60-minute semi-structured clinical interview (Ginsburg, 1997). A detailed protocol was used to guide the interview, but the interviewer followed up with clarifying questions to develop a more detailed understanding of each student's thinking. The interviews took place towards the end of the semester, and while the interview involved content introduced during the first three weeks of the course, the instructor continuously revisited this material throughout the semester. As such, students were well prepared to deal with the problems posed during the interview.

The interview protocol consisted of two main sections, with the relative data for this report taken from the first section that explored how participants interpreted and made sense of notation and logical statements in symbolic form. This section consisted of two main questions. First, the participants were asked to analyze the equivalence of various logical statements relative to $p \Longrightarrow q$ and explain whether they were equivalent or not and why. Second, the students were asked to simplify the negation of a conditional statement presented in the symbolic form $\sim(p \Longrightarrow q)$, explain their understanding of the negation, and give an example to illustrate their interpretation. In the second half of the interview, students were asked to negate several statements from various contexts and with multiple levels of complexity and explain the conditions necessary for the statement to be false. Observations from the second half were used to contextualize students' understanding, but all results reported came out of the first half of the interview.

Each interview was videotaped and transcribed. Students' responses were reviewed using a grounded theory approach (Strauss & Corbin, 1994). The initial coding pass relied on open coding in which evidence was collected to make sense of how students conceptualized logical statements and interpreted logical equivalence. After a detailed review of the videos and their accompanying transcripts, data suggested that it was the degree to which participants viewed logical statements and their associated logic tables in either a unified manner (that is, as a composite unit) or in a compartmentalized fashion that distinguished the participants' approaches to making sense of logical equivalence. After the student interviews were fully examined and reviewed, the professor of the course was then interviewed on the same questions in order to provide a comparison between the students' conceptualizations to that of an expert. We then analyzed his responses as we did for each student.

Results

While a truth table can be a very powerful organizational tool, analysis of student interviews revealed that many participants, while able to correctly construct a truth table, struggled to coordinate the various cases and to view them as a composite unit. Far from an interconnected entity, these students tended to compartmentalize truth tables, viewing each row, even each symbol, as discrete and separate pieces. Such participants showed little appreciation that the given compound statements could have four possible truth combinations, translating the notation directly or interpreting the symbolic sentences as taking on only the meaning that would most obviously make the statement true. With such a detached view, they struggled to understand the meaning of logical equivalence. Instead, they seemed to view the logic table as a mechanism to help produce and match isolated symbols, asserting that only one line of the truth table corresponding between two statements was sufficient to establish equivalence. In addition, when asked to interpret the meaning of $\sim(p \Rightarrow q)$, many participants treated the hypothesis p and conclusion q as quite separate entities. In an attempt to make sense of the negation of a conditional statement in symbolic form, every student either incorrectly distributed the negation symbol or converted $p \Rightarrow q$ to $\sim p \vee q$ in order to use de Morgan's law. Rather than interpreting $p \Rightarrow q$ as a composite unit, they attempted to manipulate each piece individually, considering each component separately.

Analysis of student reasoning in the interview tasks therefore led to the development of a two dimensional framework for characterizing the degree to which students' conceptualizations of both truth tables and implication statements are compartmentalized versus unified. The ends of each of these two dimensions are detailed in *Figure 1*.

Truth Table	
Compartmentalized	Unified
• Interpreting symbols as only taking on the meaning that makes the statement true (for example, $p \Rightarrow q$ means p is true and q is true) • Considering symbols literally or as a direct translation (for example: $\sim p \wedge q$ means p is false and q is true)	• Seeing the truth table as an organizational tool to represent the four possible cases, not to generate them • Ability to consolidate cases when strategically useful • Connecting semantic meaning to each case

Figure 1a. Compartmentalized and unified views of truth table

$p \Longrightarrow q$	
Compartmentalized	Unified
• Looking at symbols in an isolated manner without making meaning of them as a whole piece (for example, negates conditional statement by distributing negation to both *p* and *q*) • Comparing statements based on notational structure instead of meaning (for example, negation of implication must be an implication) • Tendency to invoke algebraic symbol manipulation	• Comprehending the premise and conclusion as one statement • Connecting the symbolic representation to its semantic meaning • Ability to interpret and contextualize four cases and see that three of the four satisfy the implication. In other words, connecting the understanding of the implication statement with the meaning of the truth table. • Translating symbols into meaningful, naturally linked contexts to make sense of implication

Figure 1b. Compartmentalized and unified views of $p \Longrightarrow q$ and its truth table

Based on our analysis of student responses, each student was assessed on the extent to which they exhibited a unified or compartmentalized view of truth tables and logical implications. A continuum was proposed along each of these two dimensions with students falling at different places depending on the number or degree of the various components they demonstrated. *Figure 2* depicts the findings for each student.

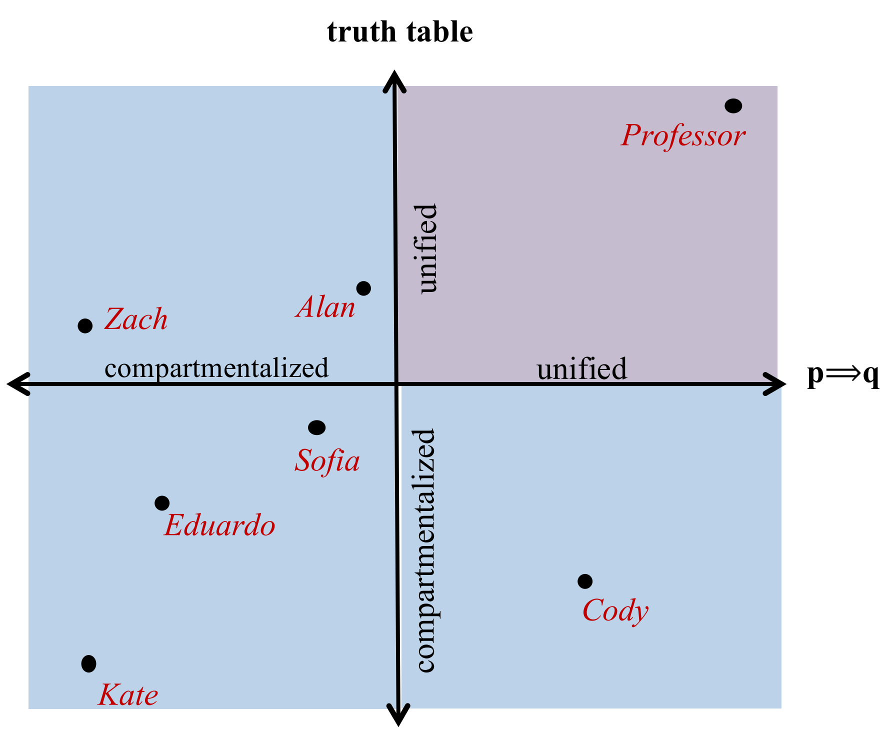


Figure 2. Framework for characterizing student thinking of $p \Longrightarrow q$ and its truth table

As illustrated in *Figure 2*, no student demonstrated a unified conception of both logical structures. While almost every student possessed a unified conception of one, the two seemed almost inversely proportional. Also, students seemed to compensate for whichever conceptualization was compartmentalized by attempting to use the other mechanism, even when it was not productive. This was in contrast to the professor who demonstrated a clearly unified conception of each logical structure. Not only did he articulate a unified understanding for both, but in his description, he was unable to disentangle the idea of an implication from its associated truth table. He quickly alternated between discussing $p \Rightarrow q$ holistically, then in parts, then in terms of its various possible truth values.

Compartmentalized view of the truth table

We begin with Kate whose understanding of the truth table, along with Cody's, is representative of other students with a strongly compartmentalized view of truth tables. As we see in the following excerpt, Kate does not interpret the statements as having multiple possible cases. Instead, she perceives each statement as having only a single true scenario and equivalence is a matter of aligning these true situations. Here she is asked to determine the equivalence of $\sim p \vee q$ and $p \Rightarrow q$,

Kate: Because 'or' just means that one or the other has to be true and we were given q is true if p is true, so this can be equivalent 'cause if we just have…yeah we can just pick one of those…yeah this wouldn't be true (writing $\sim p \wedge q$) because we don't have both of those.

When asked to explain, she replied

Kate: We have q, but we don't have not p 'cause p is not (*not p*). So we have p then q but not (*not p* and q). So that wouldn't be logically equivalent, but this would be (writing $\sim p \vee q$) because it is just one or the other or both is what this means.

When asked what it means that both are true, she said,

Kate: If you're given that this is a true statement if, p then q, then you just suppose that they are both true. For it to be a true statement, that if p then q, so p and q need both to be true.

In this last section, Kate clearly states that she interprets the implication statement as having only the one case when both p and q are true. Similarly, in the first part, it appears that she considers the symbols in the disjunction quite literally, directly translating $\sim p \vee q$ to mean p is false or q is true. Since the assumed true p and true q case from the implication aligns with the false p or true q, she asserts that the two statements are logically equivalent. This is in contrast with the conjunction ($\sim p \wedge q$) she proposes. Here the assumed true p and true q case of the implication does not match with her literal interpretation of the false p and true q of the conjunction and are therefore not logically equivalent. As this example supports, Kate' view of logical equivalence is quite fragmented, involving the comparison of isolated symbols, interpreted as superficial objects, without an appreciation of multiple cases and their necessary correspondence.

Later in the interview, after she accurately completes a truth table, her substantially compartmentalized view, as well as the subsequent consequences of such an understanding, become even more apparent. Not only does Kate's surface interpretation of compound statements cause her to incorrectly understand logical equivalence, her fragmented view of the symbols that make up the truth table prohibits her from correctly interpreting the data it represents. In the following excerpt, although she has stated earlier, based on her previous reasoning, that $\sim p \wedge q$ is not equivalent to $p \Longrightarrow q$, she reluctantly decides to change her mind when she notices the one case in implication truth table of false p with a true q resulting in true as its truth value.

Kate: I don't know if you would be able to call it logically equivalent, but it could still make it a true statement to have not p and q. So to have p be false and q be true could still create a true statement.

Intvwr: So these are logically equivalent? (pointing to p⟹q and ~p^q)

Kate: I guess you could say that. I don't know if it can be or not. I guess from the truth table I guess it could be.

Intvwr: So when you say from the truth table can you be more specific? How does that help you see that those are equivalent?

Kate: Umm… by writing it all out and then knowing, just like by memorization, that a false implies a true can be true and but true implies a false is false. Then you just see that if it's true, then it can be logically equivalent to it. So I guess if you have p then q, then it's possible for it to be equivalent to a false p and a true q. For it to still be…it's possible for p to be false and q to be true and have the statement still be true.

Intvwr: Okay, so this is…(pointing to truth table, in particular the row where p is false and q is true)?

Kate: Yeah, that's where I get it from here. These are all the possible outcomes for p and q together, so you have true – true, and false-false, and true and false, so that gives you all the different combinations of p and q and then you see what they equal true or false over them on the right side according to the different combinations over here.

As a consequence of her compartmentalized view of the symbols that make up the truth table, her correct production of a truth table in the end leads her completely astray. She does not see each of the four combinations working together to inform the possible cases of a single statement, but rather as independent pieces. Her view of these cases lacks any association and consequently she interprets each individual case as evidence of equivalence. For Kate, the truth table has no unified meaning. Produced from memory, it is simply a collection of unconnected symbols. Understanding the process of how the symbols work together is lost on her.

Partially connected view of truth table

Moving along the truth table continuum, we arrive at Zach and Alan. Their responses demonstrate the most connected view of the different cases communicated by the truth table of any of the students interviewed. They express a clear understanding that the logic symbols representing the various propositions within a given compound statement can take on different truth values and both indicate a clear understanding that equivalence requires multiple coupled cases to align. As illustrated in the following example, as Alan justifies why $\sim p \wedge q$ and $p \Longrightarrow q$ are logically equivalent, we can see that he holds a significantly more integrated view of the components that make up the truth table.

Alan: Alright, if p is something that happens, q has to happen from this statement. So then you can have either if p did not, the first thing did not happen then q could have happened or could not have happened or either one.

In this excerpt, Alan clearly outlines the three cases where the implication is true. It is evident that he recognizes that the propositions, represented by the logic symbols p and q, can take on both true and false values. It appears that as he quickly runs through the multiple cases, he is mentally checking to make sure these true cases for the implication also result in true outcomes for the disjunction (~p^q), first looking when p is true (meaning that *not p* is false) and then the cases when *not p* is true. This process demonstrates an awareness that multiple cases must correspond in order to ensure logical equivalence. Interestingly, he does not check the correspondence of all four cases, but only those that result in an overall true value. He seems to deem the true cases as the important parts of the truth table and possibly assumes that the other cases will consequently correspond. Overall, his approach, efficiently examining and comparing relevant cases without creating a list of symbols, reveals a quite connected view of the truth table. In a sense he creates a partial truth table in his head, coordinating the various cases mentally. While Alan does not elect to generate a formal table, it seems that such an exercise would have only served as a way of recording and organizing previously established verbal ideas.

Still, although Alan's method is quite effective and succinct, it appears that he has not yet fully consolidated the various cases into a single unit. The multiple cases are connected, in that he realizes that as a set they must correspond, but he still examines each possibility individually. Without seeing the three true cases as a unit, he is unable to recognize and elevate the one false case as the complement of this set, explicitly comparing and processing it along with the others.

Fully unified view of truth table

The final end point of the truth table continuum is exemplified by the professor of the class. He was interviewed subsequent to the six students to provide a point of comparison. As we will see in the following set of quotes, he demonstrated a fully unified view of the truth table. His analysis displayed an appreciation of the various cases as a single unit, comparing and manipulating them as a collective whole, instead of as a set of separate, independent possibilities. We begin with an excerpt of the professor explaining how he understands that $\sim p \vee q$ is logically equivalent to $p \Longrightarrow q$.

Dr. S: Okay, so under what conditions would p imply q? If I say, if you do x, I will do y, under what conditions would you say I've lied; I haven't kept my part of the bargain? And there is exactly one way in which that can happen. And so I see p implies q has a truth table which tells when I have told the truth and when I have been false. And that has one false in it, so that should be the same as some sort of an "or" statement.

In this selection, it is clear that he compares the four cases of the implication and disjunction simultaneously. Because the combination of three true cases and one false case matches the characteristics of an "or" statement, it makes sense that the two compound statements are equivalent. Such an examination differs considerably from that of Alan who compared the two statements case by case. Although Alan knew they were all connected and that multiple cases needed to mutually align, he did not evaluate them as a single unit. At the same time though, the professor does simply think about the statements globally. His use of the word "conditions" indicates that he is taking into consideration and analyzing the various cases as well. While he identifies the one false possibility, because he sees the four cases as a unified whole, he knows that its complement, the other three cases, must all be true and consequently align as well. Furthermore, as we see in the following selection when the professor is asked to give an example to expound on his thinking, he is able to attach semantic meaning to each of the four cases. While he does focus on the global characteristics to make his comparison, he is also able to dissect the unit and give meaning to each of the separate cases.

Dr. S: So if we went to court and you were going to build something for me, and you know, the deal was that if I paid you a thousand dollars, you would build the thing for me. If

> I didn't pay you a thousand dollars, then the contract's not broken. And if I did pay the thousand dollars and you didn't build it, then the contract is broken. If I did pay and you built it, it's satisfied. So there are three conditions in which it is satisfied and one condition in which it's not…. So if I didn't pay you, then we're okay, or if you did do it, we're okay. The second one of course is an interesting situation, I don't pay you and you build the house for me... and that's perfectly fine. The contract is satisfied. … Understanding this statement as being either true or false (pointing to $\sim p \vee q$) and it is true even when something kind of ridiculous like I don't pay you and you build the house for me.

He makes sense of the "or" statement in context, understanding that a contract is valid in the cases when payment is made or if the work is completed. It is only when the house is not built and payment is received that the conditions of the contract are violated. He even explains what it means to have a false premise and true conclusion and why this might be counterintuitive, but acceptable.

Implication statements

Implication statements provide the foundation for mathematics. As we investigate mathematical phenomenon, searching for generalities, we naturally come up with questions about whether a given set of circumstances lead to certain outcomes, writing any observations as implications statements. Consequently, an appreciation for the conditions in which a given conjecture is viable and not viable, seems fundamental to mathematical understanding. It seems reasonable that formal logic would aid in this process as it provides a structure to take a complicated statement and transform it into the simple, elementary notational representation $p \Longrightarrow q$. Unfortunately, while such a notation allows the user to hone in on the essential pieces of implication statements, making, in theory, their exploration and manipulation under different conditions much simpler, interviews indicate that students treat these symbols not as a unified whole, but as separate, disconnected entities. Instead of viewing the symbolic form as a compound statement composed of two highly associated clauses, they compartmentalize the pieces, considering them in an isolated fashion to differing degrees. This was especially evident as students were asked to interpret the negation of an implication statement in notational form $\sim(p \Longrightarrow q)$. Two distinct approaches to negation emerged, suggesting two different groupings within the participants, with exactly three students in each classification. The first set of participants (Kate, Cody, and Sofia), almost without exception, tended to treat the symbols quite superficially and incorrectly attempt to distribute the negation symbol in some way. Students from the second grouping (Eduardo, Zach, and Alan) chose to convert $p \Longrightarrow q$ into the equivalent form $\sim p \vee q$ in order to use de Morgan's law and manipulate the symbols as separate pieces. While both of these methods indicate a rather compartmentalized view of the implication statement, within these two approaches, students' interpretations still varied in the degree to which they viewed the symbols as a composite unit. In the following section we present examples of student responses from each of the two grouping along the proposed continuum, beginning with students with the least unified view.

Compartmentalized view of $p \Longrightarrow q$

Looking at students with the most heavily compartmentalized view, we begin with instantiations from the first group, those who focused primarily on superficial features when analyzing the conditional statement $p \Longrightarrow q$, turning to Sofia and Kate for examples. As we see in the following excerpt, they seem to hold the belief that equivalent statements must retain the same form. As such, an implication cannot be equivalent to a disjunction and the negation of an implication cannot be equivalent to a conjunction. This view is evident in Sofia's response as she explains why $p \Rightarrow q$ is not equivalent to $\sim q \vee p$.

Sofia: This is a conditional statement and this is, I think, just a statement. This is saying, not p or q… This one (pointing to $p \Rightarrow q$) you are saying…You are setting conditions for and then a possible outcome. So they are two… In my opinion they are two completely different things.

In this selection we see that Sofia looks to surface features of the individual symbols to compare statements instead of looking at the meaning of the implication as a whole. Consequently, later when asked to interpret $\sim(p \Rightarrow q)$, she distributes the negation without attending to the underlying meaning.

Sofia: I would just distribute the negation. So, *not p* then *not q*, is how I would interpret it. I would just distribute the negation throughout the conditional statement.

Even when asked to illustrate the negation using a contextualized example, she continues to struggle to see the implication statement as a unified whole.

Sofia: I don't know how I would write this one (referring to $\sim(p \Rightarrow q)$). But here it could be…if I don't walk the dog then he won't be happy. Here it's if I walk… I don't know how I would incorporate that without negating the whole…I can give you a real life for this one (referring to $\sim p \Rightarrow \sim q$), but just having the negation kind of hovering… I don't know how to do it in real life.

She explicitly says to make sense of $\sim(p \Rightarrow q)$ requires "negating the whole" implication statement which she is unable to. She struggles to coordinate her fragmented understanding of the symbolic form $\sim(p \Rightarrow q)$ to context, where the premise and conclusion cannot be treated separately. Instead she resorts back to her method of distribution which allows her to treat the symbols in an isolated manner and then translates them back to words.

Kate displayed even more difficulty coordinating the symbolic interpretation to a meaningful context. Although she states that $\sim(p \Rightarrow q)$ means that "p does not imply q", she continues to translate the negation symbolically as the distribution of the negation notation. When asked to create a context to illustrate her understanding, she comes up with an example where there is no meaningful relationship between the premise and conclusion. Only by providing a completely detached example can she relate her interpretation of the symbols to a real life situation.

Kate: So I guess if p is "I'm mad" and q is "I'm sad," so if you're saying if p implies q, so if I'm mad, then I'm sad. But if you're saying not all of that, you're saying if, if you're doing what I did by just distributing all of it, then I'm not mad then I'm not sad…To me it makes sense to say that if I'm not mad, then I'm not sad, because that's kind of the opposite of what I said before, if I'm mad, then I'm sad.

In this episode, it seems clear that Kate is simply substituting words piecewise for notation without any appreciation of the connection to a unified statement. Even more disconcerting, in the end, the symbolic form, instead of aiding her understanding, actually causes her to misinterpret the negation of the implication statement as she construes the meaning to be the opposite instead of refutation. Ultimately, her isolated sense of the conditional statement leads to an incorrect, reverse translation from the symbols into the words.

Looking at students from the other group who used substitution of $\sim p \vee q$ to negate the conditional statement, there were equivalent examples where they treated the symbols in a compartmentalized manner. First of all, their general method reflects a rather fragmented view. It seems the whole purpose of transforming the statement into a disjunction is to students to distribute the negation and deal with the pieces separately. Almost without exception, these students negated every implication statement throughout the entire interview using this conversion method. Even when the question was presented contextually or the students were ask to provide an explicit real life example, these students consistently changed the statement into symbols in order to make this substitution, before translating the final results back into words.

Still, while in general the approach demonstrated by these students projected a rather disjointed understanding of the symbolic structure, there were definitely different degrees depending on their ability to make sense of the negation after the substitution. As we see in the following excerpt, Zach represents the most compartmentalized view as he continues to treat the clauses as separate entities. When asked to come up with a context to illustrate his interpretation of the symbols he provides an almost absurd example, revealing no connection between the premise and conclusion.

Zach: Whatever p and q stand for, we have to know that for this statement to be true, p has to be true and q has to be false. So I don't know if you would say. Someone is tall and not tall, or something like that. It's kind of a really simple example. But…I guess it doesn't make sense, not short. So if p is tall and q is short. P is true, you are tall and they are not short, so that is true as well. So that would be a true statement, logically.

It is clear in this example that he treats p and q as completely separate entities. Instead of seeing the clauses as interrelated, one implying the other, he focuses on the fact that one is true and the other is false. He is unable to use the symbolic manipulation to aid in his understanding. In fact, the substitution seems to cause an even more disassociated view.

Partially compartmentalized view of $p \Rightarrow q$

Moving along the continuum we find examples in both groupings of students whose view of implication statements is quite connected when considered in context, but then struggle to associate this understanding to their interpretation of the symbolic form. In the first group we turn to Cody who is representative of students with inconsistencies between these two capacities. He correctly interprets $\sim(p \Rightarrow q)$ verbally as "if we had p and then, but q did not happen," further explaining his understanding that the implication "would be false...because it didn't happen," but then writes out his interpretation as $p \Rightarrow \sim q$. When asked to provide an example to illustrate his understanding, he offers up the following clarification, again demonstrating a unified consideration within context, but continuing to make the same error translating his thinking symbolically.

Cody: You are saying that not that one thing happened and the next thing happened. So like, not it's sunny outside and I went to the beach... So that would mean like it's sunny outside, but you didn't go to the beach. So that would be kind of doing what's in parenthesis, making that false. That's kind of what I interpret the not as.

It appears that Cody treats the implication statement as an integral whole, but is unable to connect this meaning to and from the associated symbols. Throughout the entire interview, regardless of the context or whether questions were presented contextually or symbolically, he repeatedly reported the correct idea verbally but described his interpretation incorrectly in terms of notation. He continued to hold on to the notion that equivalent symbolic forms must sustain a consistent structure, treating the notation in a more compartmentalized manner. Consequently, he was unable to manipulate or make use of the decontextualized symbols in order to guide his understanding. As such, the logical notation is not a tool for Cody, as any support the more abstract, simplified presentation could offer is lost on him. Nonetheless, he does seem to hold a fairly unified view of the implication statement in contextual situations.

In the other grouping, we turn to Alan as representative of the middle level along the continuum in terms of a unified view of the implication statement. Like the other students in this classification, he approached the negation of almost every problem (4 of 5 questions) by translating the implication to the equivalent disjunction in order to be able to distribute the negation sign. While this approach, in general, would indicate a rather disjointed view of $p \Rightarrow q$, Alan's understanding differs though from the other's using the same technique in that he is able to connect his final symbolic result back to the original implication in a unified manner through context. In the following excerpt he has just provided an example to illustrate the negation of $p \Rightarrow q$ and he is explaining how his final result of $p \wedge \sim q$, arrived at through symbolic manipulation, is a sensible outcome.

Alan: Okay, so if I studied for the test, then I got an A on the test. If I didn't get an A on the test then I wouldn't have studied, because studying would imply that I got the A.

Through this example we first see that Alan provides a coherent context to illustrate the meaning of an implication, allowing him to effectively link his thinking. Unlike Zach who chose a completely dissociated premise and conclusion, Alan's example is well connected. This provides evidence that he is not simply translating the symbols independently, but sees the relationship between the two clauses, seeming to hold a more unified view of the implication. In addition, while he is unable to initially manipulate the symbolic form as a unified unit, he does connect his notational translation back to the original problem. Through his tone as he reads through the result, it seems clear that he sees how such an outcome is the reasonable negation of $p \Rightarrow q$. Holding a more integrated view of the implication statement allows him to make sense of the symbolic representation.

Fully unified view of $p \Rightarrow q$

As was the case with the truth table, no student demonstrated a fully unified view of $p \Rightarrow q$, with the only such example coming from the professor. Throughout the interviews each of the six students repeatedly attempted to manipulate the different pieces of the symbolic form of the implication statement independently. In addition, not only did the students treat the implication statement in a piecewise fashion, none of the participants expressed any consideration of multiple cases when negating $p \Rightarrow q$. This is in contrast to the professor's conceptualization, where both the implication and its associated truth table are completely intertwined. As was seen in the previous examination of the professor's understanding of the truth table, not only does he view the implication as a unified whole, he interpreted it through its four associated cases. Having consolidated these possibilities into a single unit, he is able to see negation of the statement as the one false case he has identified. Consequently, the idea of a fragmented view of $p \Rightarrow q$ makes no sense, because his interpretation of the implication statement is interwoven with the four cases. Furthermore, the contextual example that the professor attaches to the implication statement reinforces his understanding of $p \Rightarrow q$ as a unified object. The situation of a binding contract depicts the implication statement as a single unit, consisting of a particular premise and conclusion that may or may not happen. Negation of the implication means breaking the agreement and can only be understood by thinking of what combination of initial and resulting conditions would invalidate the contract. As such, it seems clear that the professor sees the contract as a composite unit, unable to view p and q as completely disconnected. In the end, it is his coordination between the two logical mechanisms along with a well-chosen context that supports the professor's unified view of the implication statement.

Conclusion

Mathematical notation is a powerful tool. Representing concepts as concrete symbols provides a medium to communicate and reason about abstract mathematical ideas. An inherent feature of mathematical notation is that concepts are represented using discrete characters. As such, individual parts and pieces can be operated on and treated independent of any overarching concept or as objects without consideration of their underlying semantic meaning. While such an attribute is an affordance for mathematicians, reducing the associated cognitive demand by breaking down complicated concepts and allowing for manipulation without constant attention to what it references, it can also be a limitation for students who are only beginning to familiarize themselves with the notation, hindering their appreciation of the concept as an integral whole and enabling them to acquire a compartmentalized view. It is therefore necessary, when working with mathematical notation, to develop the capacity to flexibly contextualize and decontextualize the symbols; to simultaneously see the notation representing a unified concept as well as its various constituent pieces. This is true of symbols associated with all areas of mathematics, including logic. Focusing on the degree to which students demonstrated a conceptualization of truth tables and implication statements as composite units proved a beneficial perspective to characterize the thinking of students. It provided a framework to interpret and make sense of students' understanding of these symbolic tools and their associated concepts. Results include specific characterizations of what a compartmentalized and unified view of truth tables and conditional statements looks like, operationalizing both ends of the continuum.

The findings of this study suggest that without an appreciation of truth tables as a composite unit, students struggle to understand logical equivalence. Seeing only fragmented pieces, equivalence becomes a matching exercise of one perceived representative case or multiple independent cases, but lacking a full appreciation of their meaning. Without a view of the various truth value permutations as a single element, students are unable to understand how structurally different statements can be equivalent and when and why a transformation from one form to another is allowed.

Similarly, interview data indicates that students who fail to consider implication statements as a single meaningful unit, struggle to understand negation. Seeing each symbol as a separate entity, many participants carried out incorrect operations, often applying notions of the distributive property from algebra, without an appreciation of any underlying meaning. Even students who correctly manipulated the symbolic forms and arrived at a valid notational representation of a negated implication statement, often did not demonstrate seeing a connection between the premise and the conclusion. Converting to and from any contextualization of the statement appeared to be simply substitution. These students tended to view syntactical forms of implication statements as purely symbols, divorcing their understanding of the logical notation from their personal experience reasoning through problems. Consequently, the symbolic form did not appear to be an aid in supporting their understanding.

Finally, not only did the students possess a compartmentalized view of both truth tables and logical statements, their coordination between the two was also disjointed. They treated the two mechanisms as separate and developed a more unified view of one, relative to the other. Consequently, students tended to compensate for whichever conceptualization was compartmentalized by attempting to use the other mechanism, even when it was not productive. This was in contrast to the professor. The unified view he held of both of these tools converged. He could not understand the negation of a unified implication statement, without seeing its associated truth table as a composite unit.

References

Barnard, T. (1995). The impact of meaning on students' ability to negate statements. In *PME conference* (Vol. 2, pp. 2-3). The program committee of the 19th PME conference.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18(3), 293-328.

Common Core State Standards Initiative. (2010). Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/math.

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. Schoenfeld & J. Kaput (Eds.), *Research in Collegiate Mathematics Education IV - Issues in Mathematics Education* (Vol. IV): CB MS.

Epp, S. S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly*, *110*(10), 886-899.

Ginsburg, H. (1997). *Entering the child's mind*. Cambridge, UK; Cambridge University Press.

Herbst, P. G. (2002). Establishing a custom of proving in American school geometry: Evolution of the two-column proof in the early twentieth century. *Educational Studies in Mathematics*, *49*(3), 283-312.

Lamon, S. J. (1996). The development of unitizing: Its role in children's partitioning strategies. *Journal for Research in Mathematics Education*, 170-193.

Knuth, E. J. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, Vol. 33, No. 5 (Nov., 2002), 379-405.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, 27(3), 249-266.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Rav, Y. (1999). Why do we prove theorems?. *Philosophia mathematica*,*7*(1), 5-41.

Ross, K. (1998). Doing and proving: The place of algorithms and proof in school mathematics. *American Mathematical Monthly*, 3, 252-255.

Selden, J. and A. Selden (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics,* 29, 123-151.

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification—the case of algebra. In *Learning mathematics* (pp. 87-124). Springer Netherlands.

Sfard, Anna. "The development of algebra: Confronting historical and psychological perspectives." *The Journal of Mathematical Behavior* 14.1 (1995): 15-39.

Steffe, L. P. (1983). Children's algorithms as schemes. *Educational Studies in Mathematics*, 14(2), 109-125.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273 – 285). Thousand Oaks: Sage Publications.

Wheatley, G. H., & Reynolds, A. (1996). The construction of abstract units in geometric and numeric settings. *Educational Studies in Mathematics*, *30*(1), 67-83.

TWO METAPHORS FOR REALISTIC MATHEMATICS EDUCATION DESIGN HEURISTICS: IMPLICATIONS FOR DOCUMENTING STUDENT LEARNING

Estrella Johnson
Virginia Tech

The primary goal of this work is to articulate a theoretical foundation based on Realistic Mathematics Education (RME) that can support the analysis of student learning. I first describe two RME design heuristics, guided reinvention and emergent models, and explicate each of these heuristics in terms of related theoretical constructs. I then consider how the RME design heuristics could inform how one conceptualizes of and documents student learning (where learning is viewed in terms of the creation of a new mathematical reality). To do so, I draw on two metaphors for learning and, by considering the design heuristics in light of these two perspectives, I propose two ways to conceive of "new mathematical reality" and discuss what could be considered as evidence for student learning.

Key Words: Realistic Mathematics Education, Learning, Analytic Methods

Realistic Mathematics Education (RME) is an instructional design theory used to inform the development of inquiry-oriented curriculum. The emergence of such instructional approaches creates a need to investigate student learning in these contexts. However, as it was designed to be an instructional design theory, the current formulations of RME are not articulated in a way that readily supports investigations of student learning. Part of the difficulty in using the current formulations of RME to investigate student learning is due to variations in the ways that the RME design heuristics (specifically *guided reinvention* and *emergent models*) are discussed in the research literature. For instance, both the guided reinvention and emergent models design heuristics are described in terms of the creation of a *new mathematical reality*. What exactly a new mathematical reality is, however, remains unclear. At times, the creation of a new mathematical reality is discussed as being equivalent to activity (Rasmussen, Zandieh, King, & Teppo, 2005). Other times the creation of a new mathematical reality is discussed in terms of object reification. As a result, efforts to document the creation of a new mathematical reality are not supported by a clear theoretical foundation. In order to articulate RME in a way that supports analytic techniques for documenting student learning, I will first describe two RME design heuristics, guided reinvention and emergent models, and explicate each of these heuristics in terms of related theoretical constructs. I will then draw on two metaphors for learning, and propose ways in which the RME design heuristics can inform the analysis of student learning.

Guided Reinvention and Emergent Models

RME is grounded in the belief that mathematics is "first and foremost an activity, a *human* activity" (Gravemeijer & Terwel, 2000, p. 780). Accordingly, Freudenthal argued that mathematics education should "take its point of departure primarily in *mathematics as an activity,* and not in mathematics as a ready-made-system" (Gravemeijer & Doorman, 1999, p. 116). Within RME there are a number of heuristics that are meant to guide the design of instruction that supports students in developing formal mathematics by engaging them in mathematical activity. For both the *guided reinvention* and the *emergent model* heuristics, the nature of this activity becomes more general as the instructional sequence unfolds.

Guided Reinvention

Guided reinvention is often characterized in terms of the nature of the learning process, where the goal is for "learners to come to regard the knowledge they acquire as their own, personal knowledge, knowledge for which they themselves are responsible" (Gravemeijer & Terwel, 2000, p. 786). To achieve this goal, mathematics instruction is designed around mathematical actives intended to expand the students' common sense. As Gravemeijer (1999) describes, "what is aimed for is a process of gradual growth in which formal mathematics comes to the fore as a natural extension of the student's experiential reality" (p. 156).

The student's experiential reality includes what a student can access on a commonsensical level. As Freudenthal (1991) discusses, ""Real' is not intended here to be understood ontologically (whatever ontology may mean), therefore neither metaphysically (Plato) nor physically (Aristotle); not even, I would even say, psychologically, but instead commonsensically" (p. 17). Accordingly, the problem context that serves as the basis of the reinvention process need not be "real" in the sense that the students would access such scenarios in their everyday life. Instead, the students only need to be able to access the problem context on an intuitive level. In this way, the movements of a magic carpet may provide a context for the reinvention of the formal mathematics of linear algebra (Wawro, Sweeney, & Rabin, 2011).

Within an experientially real context, the reinvention process progresses through a series of instructional tasks that promote mathematizing the problem context. This activity of mathematizing, "which stands for organizing from a mathematical perspective" (Gravemeijer & Doorman, 1999, p. 116), is viewed as the mechanism through which students reinvent the mathematics. In fact, Garvemeiher (1999) asserts that it is "via a process of progressive mathematization, the students should be given the opportunity to reinvent mathematics (p. 158). Because progressive mathematizing has been described in the RME literature as the primary mechanism supporting guided reinvention, my discussion of the connections between guided reinvention and student learning will be focus on that mechanism. Progressive mathematizing is typically described as consisting of cycles of horizontal and vertical mathematizing.

Initially, as students mathematize their own experiential reality, they are engaging in horizontal mathematizing. Horizontal mathematizing could include activities such as translating, describing, and organizing aspects of problem context into mathematical terms (Gravemeijer & Doorman, 1999). It is the nature of the artifact of the activity that provides an indication that horizontal activity has taken place. The artifacts of horizontal mathematizing, which may include inscriptions, symbols, and procedures, are used by the students to "express, support, and communicate ideas that were more or less already familiar" (Rasmussen et al., 2005, p. 164).

While horizontal mathematizing is a crucial step in the reinvention process, reinvention "demands that the students mathematize their own mathematical activity as well" (Gravemeijer & Doorman, 1999, p. 116-177). As students mathematize their own activity, they engage in vertical mathematizing. Vertical mathematizing activities are characterized by the nature of the subject matter being mathematized, where it is the student's mathematical activity that is now the subject of vertical mathematizing. This distinction between the context of the mathematizing activities, with horizontal mathematizing acting on familiar problem contexts and vertical mathematizing acting on mathematical activity, leads to differences in the artifacts resulting from these two forms of mathematizing. Instead of using the artifacts to describe already familiar situations and ideas (as with horizontal mathematizing), artifacts of vertical mathematizing can be used by students in more general settings to describe, express, and create previously unfamiliar mathematical ideas. In this way, vertical mathematizing expands what is experientially real for the students by establishing a new mathematical reality. Examples of such activity include generalizing, defining, and algorithmatizing (Rasmussen et al., 2005).

An example of progressive mathematizing in abstract algebra. To illustrate such a reinvention process, and the cycles of progressive mathematizing, consider an example from an RME inspired, abstract algebra curriculum – *Teaching Abstract Algebra for Understanding* (TAAFU). The TAAFU curriculum was designed to be used in proof based, introductory group theory courses at the undergraduate level. The TAAFU curriculum includes three main instructional units: groups and subgroups, isomorphism, and quotient groups. Each of these three units begins with a reinvention phase, where the students work on a sequence of tasks designed to help them develop and formalize a concept by drawing on their prior knowledge and informal strategies. The end product of the reinvention phase is a formal definition and a collection of conjectures (for a detailed description of TAAFU see Larsen, Johnson, and Weber, 2013).

The quotient group unit is launched in the context of the symmetries of a square. By this point in the curriculum the students have worked extensively with symmetry groups as they reinvented the concepts of group and isomorphism (see Larsen, 2013). As a result, the group of symmetries of a square (and the associated operation table) is experientially real to the students, in that this group is accessible on an intuitive level. Also available within the students' experiential reality is the behavior of the even and odd integers, specifically the pattern that *even + even = even*, *odd + even = odd*, *even + odd = odd,* and *odd + odd = even*. The students are asked if they can find anything like the evens and odds in the symmetries of a square. This task represents a horizontal mathematizing activity because the students are being asked to mathematize two already familiar contexts, the symmetries of a square and the even/odd pattern.

One possible artifact of this horizontal mathematizing may be a partition of the symmetries of square. In Figure 1 we see a student's partitioning in which the symmetries of a square are divided into the rotational symmetries and the flip symmetries. In the TAAFU curriculum, the students are then asked to further mathematize their activity (and the associated artifact) by determining if this partition satisfies the definition of a group. Because the students are now mathematizing their own mathematical activity, determining if such a partition forms a group is an example of vertical mathematizing. In the course of this activity, the students determine that this partition could be viewed as a special type of a group – one in which the two elements of the group are subsets and the operation between any two subsets is determined by combining each element of one subset with each element of the other subset. In this way, the artifact of this vertical mathematizing activity is a new way to think about partitions and a new kind of group.

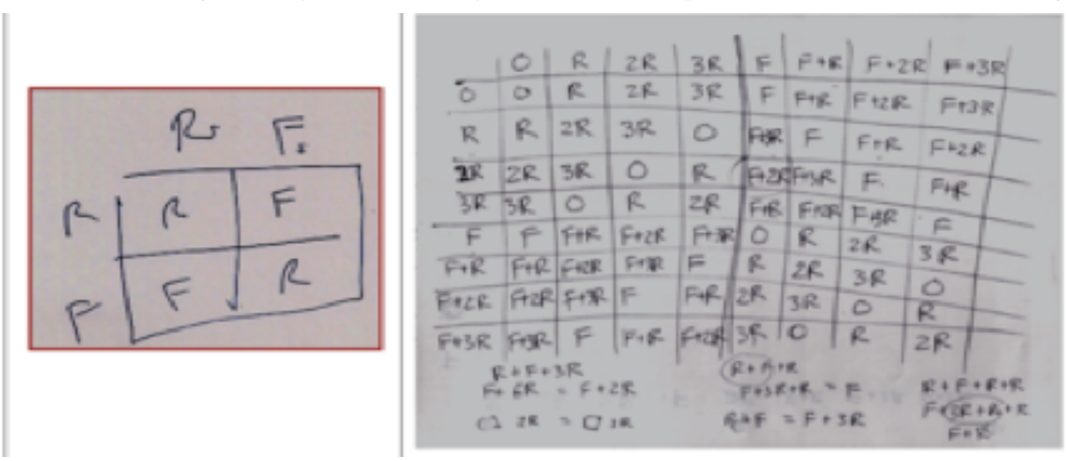


Figure 1. An even/odd partition of the symmetries of a square

The TAAFU curriculum then asks the students if they can make a larger group by breaking the symmetries of a square into four subsets. Notice that this task is not explicitly asking the students to mathematize their previous activity (i.e., their new example of group). Instead, this task is asked from the perspective that the experientially real problem context has been expanded to include this new notion of a partition forming a group. This reflects a perceived shift in the mathematical reality, where this new type of group (an artifact of vertical mathematizing) is now accessible to the students. Therefore, the task posed to the students (to make a larger group by partitioning into more subsets) is an example of horizontal mathematizing. The artifact of this task is the symmetries of the square portioned into four subsets. The students are then again asked to engage in vertical mathematizing, as they determine if this four-element partition forms a group. This round of vertical mathematizing results in a more generalized view of partitions that form groups, and ultimately a working definition of (and means for constructing) a quotient group - a group of subsets under the operation of "set multiplication". In this way, the TAAFU curriculum guides students in reinventing the concept of quotient group, through a process of progressive mathematizing.

Emergent Models

One approach to supporting students' reinvention of mathematics is to design starting point tasks that can elicit informal student strategies that anticipate more formal mathematics. The RME emergent models instructional design heuristic (Gravemeijer, 1999) is meant to support this approach. In the emergent model heuristic, informal and intuitive models of students' mathematical activity transition to models for more formal activity. A model is considered a *model-of* when an expert observer can describe the students' activity in terms of formal mathematics that is the target of the instructional sequence (Larsen & Lockwood, 2013). The model later evolves into a *model for* more formal activity. The model is considered to be a *model-for* when students can use the model to support more general reasoning in new situations.

In describing the progression from a model-of informal activity to a model-for more formal mathematics, Gravemeijer (1999) discusses four layers of activity. Initially student activity is restricted to the *task setting*, where their work is dependent on their understanding of the problem setting. *Referential activity* develops as students construct models that refer to their work in the task setting. *General activity* is reached when these models are no longer tied to the task setting. Finally, *formal activity* no longer relies on models. In regards to these four levels of activity, the shift from *model-of* to *model-for* is said to occur as students shift from *referential activity* to *general activity*. It is during this transition from referential to general activity that "the model becomes an entity in its own right and serves more as a means for mathematical reasoning than as a way to symbolize mathematical activity" (Gravemeijer, 1999, p. 164).

During the instructional sequence the "model manifests itself in various symbolic representations" (Gravemeijer, 1999, p. 170). The *chain of signification* construct provides one way to describe changes in the symbolic representation of the model during an instructional sequence, and ultimately the evolution of the global model. Central to the chain of signification construct is the idea of a sign, which is made up of a signifier (a name or symbol) and the signified (that which the signifier is referencing, such as the students' activity). A "chain of signification" occurs as students' previous signs become the signified in subsequent signs. When this happens, it is said that the initial sign has slid under the subsequent sign. These local shifts in the *form* of the emerging model support the evolution of the global model in a number of ways. As Gravemeijer (1999) notes, "the chain of signification is in a sense the counterpart, on a more specific level, of what the model is on a more general level" (p. 175). When one sign slides under, the new symbol efficiently encapsulates the students' previous activity. In this way, the new sign serves to condense the earlier rounds of activity - placing the most general activity at the forefront of the chain while still allowing students to access their earlier activity if needed. Additionally, as the instructional sequence progresses, the constant revision of the signs ensures that the current sign is the most useful for the students' current activity.

While a chain of signification looks at the development of the model on a local scale by focusing on the form of the model, the transition from a *record-of* to a *tool-for* serves as a way to understand the development of the model on a local scale by focusing on the function of the model. As described by Larsen (2004), an inscription representing students' mathematical activity transitions from a *record-of* to a *tool-for* when the students use the notational record to achieve subsequent mathematical goals. Therefore, instead of focusing on the relationships between the students' emerging symbols and notations (as with chains of signification), the record-of/tool-for construct focuses on changes in how the emerging symbols and notations are used. These local shifts in the *function* of the emerging model support the evolution of the global model in a number of ways. For instance, a local of/for shift may indicate that one aspect of the student's activity has become available to the student for more formal reasoning. The availability of this new tool reflects that certain aspects, or certain representations, of the global model are beginning to transition to a model-for more formal activity.

Chains of signification and the record-of/tool-for construct provide lenses for describing local shifts in the various symbolic representations of the global model. The former attends to changes in the form of these symbolic representations, and the later attends to changes in the function of these symbolic representations. These local changes also support each other. Changes in the form of the representation provide students with more powerful inscriptions that better meet the needs of their current activity. As a result, these new inscriptions are more useful as tools. Additionally, as students change the function of the inscriptions to achieve new goals, they may adopt more efficient forms of the representations that highlight aspects that are especially useful. Therefore, the chains of signification and the record-of/tool-for constructs are reflexively related and work together to support changes in the global model.

An example of emergent models in abstract algebra. In the first unit of the TAAFU curriculum, the students reinvent the group concept by investigating the symmetries of an equilateral triangle (see Larsen, 2013). Here the model is considered to be the algebraic structure of this particular group (the symmetries of an equilateral triangle). The students begin the unit by physically manipulating an equilateral triangle. Within this initial task setting, students identify, describe, and symbolize all of the symmetries of an equilateral triangle. The students then engage in referential activity when they being to manipulate the symbols that represent the symmetries of the triangle. This referential activity includes working with these symbolic representations in order to create a method for calculating combinations of symmetries. It is this referential student activity that an expert observer can describe in terms of the algebraic structure of this group. For instance, as students calculate certain combinations, they may regroup pairs of symmetries (implicitly using the associativity property) and notice that certain pairs of symmetries undo each other (implicitly using inverses). In this way, the structure of this group can be seen as a *model-of* the students' intuitive and informal activity. As the instructional sequence unfolds, the students' activity progresses to generalized activity as they use the algebraic structure of this group of symmetries to analyze other systems (e.g., the integers under addition) and ultimately develop a formal definition of group. At this point, the concept is considered to be a *model-for,* as the students can use the concept to support more general reasoning in new situations.

In this example the model (i.e., the algebraic structure of the group of symmetries of an equilateral triangle) undergoes a series of local changes in various symbolic representations of the students' activity. These symbolic representations include the list of symmetries, an operation table, and a set of rules for manipulating symbols. The model comprises the collection of these representations, along with the connections between them. As a result, the development of this global model is reflexively related to 1) the development of more powerful forms of the model (as described by the chain of signification construct), and 2) the students' increasing ability to reason with the various forms of the model (as described by the record-of/tool-for transition).

Chain of signification. Initially, the students begin the group unit by physically moving a triangle in order to identify the six symmetries of an equilateral triangle. The students are then asked to represent these six symmetries with a diagram, a written description, and a symbol (see Figure 2). This set of inscriptions can be thought of as a signifier that signifies the students' activity of manipulating the triangle. The students are then asked to generate a new set of symbols, this time representing each symmetry in terms of a vertical flip, F, and a 120^{o} clockwise rotation, R. This new set of symbols represents the next step in the chain of signification, with the earlier sign sliding under this subsequent sign. The original sign, which was composed of both the students' initial signifier (i.e., their initial inscriptions) and the original signified activity (i.e., physically manipulating the triangle), is now signified by this new set of symbols. So as the chain builds, students no longer need to directly consider their original activity of manipulating the triangle. For example, when working with symbols expressed in terms of F and R, students may no longer need to keep in mind that they refer to motions of a triangle.

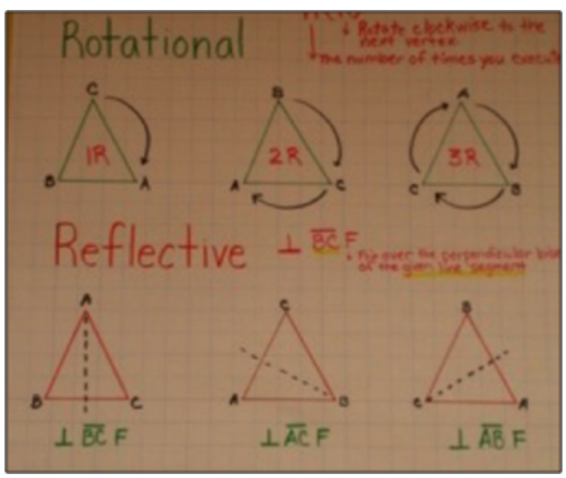


Figure 2. Diagrams and initial symbols for the symmetries of an equilateral triangle

This new set of symbols (in terms of two generators F and R) supports the students in developing a set of rules for calculating the combination of any two symmetries. For instance, when combining a vertical flip (F) and a flip over one of the diagonals ($F + R$) the symbolic expression $F + (F + R)$ invites regrouping the F's and then ignoring them because performing a flip twice is the same as doing nothing[1]. Therefore, it is in the students' use of these more

[1] With the students original symbols (shown in Figure 2) this combination would be represented by $(\perp\overline{BC}\ F)$ $(\perp\overline{AB}\ F)$, which does not suggest the possibility of rule-based calculations the same

efficient symbols (in terms of F and R) that an expert can begin to describe the students' activity in terms of the model. Specifically, as students begin to develop rules for calculating combinations of these symmetries, an expert would be able to recognize the algebraic structure of the group of symmetries as a *model-of* the students' activity. As a result, the sliding under of the students' initial inscriptions helped to support the development of the students' emerging model by 1) providing them with powerful inscriptions, and 2) transitioning their activity from the initial task setting to referential activity.

Tool-of/record-for. Once the students develop a common set of symbols using F and R, they are asked to determine the result of the combination of any two symmetries. As a way to organize the 36 different combinations, some students choose to record their results in an operation table. In this way, an operation table can initially emerge as a *record-of* the student's activity. This *record-of* can later be used by students as a *tool-for* subsequent mathematical activity. For instance, partial operation tables can be used by students to argue that the identity element of a group must be unique. This was the case during a teaching experiment that supported the development of the TAAFU curriculum. In this teaching experiment, a student worked to show that the identity element of a group must be unique. This student employed a partial operation table, with arbitrary elements, in order to construct a proof (see Figure 3). As discussed by Weber and Larsen (2008), the student's modification of the operation table, including the use of arbitrary elements and only including aspects of the table that were needed to support her reasoning, "suggests that she was using the table as a tool to support her reasoning and not merely as a crutch for recalling the steps of a procedure" (p. 148).

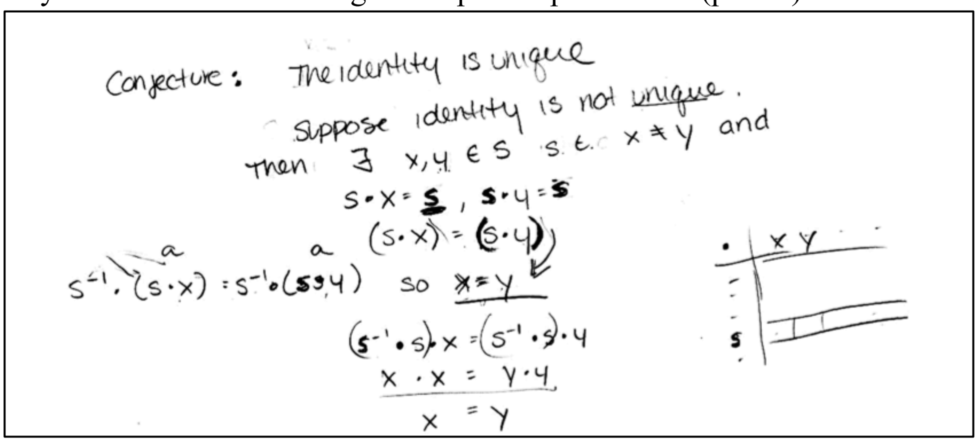


Figure 3. Operation table as s tool-for

In this example, the student was able to use an operation table to prove that the identity of a group must be unique. This represents the operation table shifting from a record-of the student's activity to tool-for further mathematizing. This shift in the function of the operation table, from an inscription (of one aspect) of the model to an instrument that can used for justification, can be seen as a local change that supports the global of/for transition of the model in two important ways. First, the operation table represents one aspect of the global model. Therefore, this local of/for shift in the function of the operation table reflects that one aspect of the global model is now available to the students for more formal reasoning. Second, the student was able to leverage the operation table as a tool as she reasoned about the formulation of the identity property. In this way, the operation table served as a tool for supporting the development of

---

way that F +(F + R) does (See Larsen (2009, 2013) for a more detailed discussion of the use of compound symbols).

another aspect of the global model – the axioms that characterize the algebraic structure of this group. Therefore, this local shift of the operation table from a record-of activity to a tool-for further mathematizing aided in the development of other aspects of the global model, and reflects an increasing ability to reason with the various forms of the model.

Framing RME Design Heuristics as Lenses on Student Learning: Two Metaphors

Captured within both the guided reinvention and the emergent model heuristics is the duality of engaging in more generalized activity and developing mathematical concepts. By teasing apart these two aspects, two lenses for describing the purpose of these RME design heuristics come into focus. One lens, which considers the guided reinvention and emergent model heuristics in terms of more generalized student activity, places the emphasis on instruction that promotes "socially and culturally situated mathematical practices" (Rasmussen et al., 2005, p. 55). The other lens, which considers the guided reinvention and emergent model heuristics in terms of concept development, places the emphasis on instruction that supports the reification of student activity. In the following section I draw on Sfard's (1998) two metaphors for learning, the participation metaphor and the acquisition metaphor, to provide insight into how the design heuristics support student learning. Additionally, by considering the implication of these two metaphors, I will present two conceptualizations of the notion of a "new mathematical reality".

Participation Metaphor and the Creation of New Mathematical Realities

Sfard (1998) describes the *participation metaphor* for learning as a view in which "learning" is synonymous with becoming a participant in a community, and "knowledge" is synonymous with aspects of practice/discourse/activity (p. 7). With this view, the emphasis is placed on what the student is *doing*, and the context in which that practice is taking place (as opposed to emphasizing the mental constructs the students have). This emphasis on the students' activity offers a lens to describe guided reinvention and emergent models in terms of more generalized activity that the heuristics are intended to support.

With the guided reinvention heuristic, instruction can be designed with the purpose of supporting student activity through progressive mathematizing. During the process of progressive mathematizing, the students' activity shifts repeatedly from horizontal to vertical mathematizing. This shift in the type of mathematizing corresponds to a shift in the generality of the student activity. Initially, horizontal mathematizing is limited to the specific problem context. As students transition to vertical mathematizing, this specific problem context is no longer the focus of the activity, rather the students mathematize their own mathematical activity to support their reasoning in a different or more general situation. Similarly, within the emergent models heuristic, there is an intention to progress students from activity situated within a specific task context to referential, general, and formal activity. In particular, the *model-of/model-for* transition is linked to a shift in the students' activity from referential (where their activity references aspects of the original task setting) to general (where the students activity is no longer tied to the original task setting). The activity that supports the transition from a *model-of* to a *model-for* is an especially significant example of vertical mathematizing. When the students are engaged in referential activity, the model emerges as a result of the students mathematizing the problem context (i.e., horizontal mathematizing). As the students move into general activity, they begin to mathematize aspects of their emerging model. In this way the transition between referential and general activity can be interpreted as the result of vertical mathematizing.

By focusing on the student activity, we get a framing of the RME design heuristics that is consistent with a participation metaphor for learning. With this framing, the process of progressive mathematizing (guided reinvention) and the progression thorough more general layers of activity (emergent models) align with Rasmussen et al.'s (2005) notion of *advancing mathematical activity* – where advancing mathematical activity is understood as "acts of participation in different mathematical practices" (p. 53). From this perspective, "participation in these practices, and changes in these practices, is synonymous with learning" (p. 55). Continuing with the participation metaphor, one could ask what it means for student activity (i.e., learning) to support the development of a new mathematical reality. I propose that, from a participation perspective on learning, the creation of a new mathematical reality can be understood as the creation of a new context for further activity, and new ways for students to participate in that context.

Acquisition Metaphor and the Creation of New Mathematical Realities

With the *acquisition metaphor*, learning is viewed as the acquisition of knowledge and concepts. This perspective "makes us think about the human mind as a container to be filled with certain materials and about the learner as becoming an owner of these materials" (Sfard, 1998, p. 5). This perspective places the emphasis on concept development, where "concepts are to be understood as basic units of knowledge that can be accumulated, gradually refined, and combined to form ever richer cognitive structures" (p. 5). This perspective comes to the forefront when the guided reinvention and emergent models heuristics are framed in terms of reification (e.g., Gravemeijer, 1999; Gravemeijer & Doorman, 1999). As described by Sfard (1991) mathematical objects (such as numbers and functions) historically developed through a recurring pattern of reification, in which "various processes had to be converted into compact static wholes" (p. 16). Similarly, one can conceive of the guided reinvention and emergent models heuristics as processes through which student activity becomes reified into mathematical objects. This emphasis on reification offers a lens to describe these two design heuristics in terms of the development of the concept, where aspects of the students' mathematical activity become reified as they engage in more general activity.

With guided reinvention, mathematical concepts develop as a result of horizontal and vertical mathematizing. By engaging in horizontal mathematizing, the students translate aspects of their mathematical reality into mathematical terms. The artifacts of horizontal mathematizing may include inscriptions, symbols, and procedures that represent aspects of an already familiar problem context. During vertical mathematizing, it is the students' own horizontal mathematizing (and resulting representations/artifacts) that are mathematized. In this way, the students' activity becomes the subject matter for subsequent mathematical activity. As described by Freudenthal (1971), "the activity on one level is subjected to analysis on the next, the operational matter on one level becomes a subject matter on the next level" (p. 417). It is this shift, from "operational" to "subject matter", that Gravemeijer and Terwel (2000) state is related to reification, where this shift reflects that aspects of the students' activity have evolved "into entities of their own" (p. 787). Similarly, the shift from model-of to model-for is related to the process of reification (Gravemeijer, 1999). As students shift from referential activity to general activity "the model becomes an entity in its own right and serves more as a means of mathematical reasoning than as a way to symbolize mathematical activity grounded in particular settings" (p. 164). Therefore, the model – which Gravemeijer (1999) describes as "an overarching concept" (p. 170) – transitions from an artifact of the students' mathematical activity to a mathematical object independent of the students' original activity. In this way the shift from model-of to model-for, which can be understood to be the incorporation of this new object into the students' experiential reality, reflects the creation of a new mathematical reality.

By focusing on the reification of student activity, and therefore on concept development, we get a framing of the RME design heuristics that is consistent with an acquisition metaphor for learning. From an acquisition perspective, learning not only supports the creation of a new mathematical reality (as it did with the participation metaphor), learning can be viewed as synonymous with the creation of a new mathematical reality. I propose that, from an acquisition perspective on learning, the creation of a new mathematical reality can be conceptualized as the incorporation of new mathematical objects into the students' experiential reality. These new mathematical objects can be understood as concepts that form "ever richer cognitive structures" (Sfard, 1998, p. 5), and the fact that they become incorporated into the students' experiential reality reflects that the students are able to access these concepts on an intuitive level.

Implications for Analyzing Student Learning

The participation and acquisition metaphors offer two perspectives on student learning – one primarily focusing on the activity of the students and the other primarily focusing on the development of the mathematical concepts. These two perspectives can provide theoretical support for analytic techniques designed to document student learning, especially in classrooms with RME inspired curricula materials. For instance, if the curricular materials were designed to encourage student learning by way of an emergent models transition (either because the engagement in such activity *is* learning or because engaging in such activity *supports* learning by developing a mathematical concept), then attempts to document student learning could explicitly draw on the theoretical constructs related to such a transition. In this section I will consider implications for analyzing student learning in cases where the instructional design is consistent with these RME heuristics (guided reinvention and/or emergent models). Specifically, I will consider what could count as evidence of student learning by considering examples from the TAAFU curriculum.

Evidence for Student Learning from a Participation Perspective

In order to document student learning from a participation perspective, the focus of analysis needs to be placed on 1) the students' participation in mathematical practices, and 2) changes in the mathematical practices of the students. The RME design heuristics provide a lens for looking at: the nature of the practices that students engage in (mathematizing), the trajectory of the students' activity in terms of generality (progressive mathematizing and layers of generality), and changes in the context and the ways students participate in that new context (new mathematical realities).

Because student learning is taken to be synonymous with participation in practices (and changes in those practices) documenting the mathematical activity of the students is a necessary for document student learning. Central to the guided reinvention heuristics is *mathematizing.* While a complete taxonomy of such practices has not been compiled, there are several examples of mathematizing practices in the literature. These include: translating, describing, organizing, symbolizing, algorithmatizing, defining, and generalizing (Gravemeijer & Doorman, 1999; Rasmussen et al., 2005; Zandieh & Rasmussen, 2010). For instance, in the quotient group unit of the TAAFU curriculum, tasks are designed to support students in *organizing* the operation table of the symmetries of a square into an even/odd pattern, *proving* that some partitions of groups themselves form groups, and *defining* a new category of group. When taking a participation perspective on learning, the documentation of student engagement in such activity provides evidence of learning.

Analysis can go even further if the instructional sequence was designed in accordance with the guided reinvention or emergent models heuristics. Both design heuristics, as used by the curriculum developer, inform a hypothesized trajectory that the instructional sequence is designed to support. These hypothesized trajectories can serve as a guide for analyzing the progression of the students' participation in mathematical activity. With the guided reinvention design heuristic, the instructional sequences are designed to support hypothesized trajectories of progressive mathematizing. With emergent models, the instructional sequences are designed to support hypothesized trajectories of increasingly general activity. Therefore, with either design heuristic, student learning can be documented in relation to a hypothesized trajectory of student activity. For instance, in the group unit of the TAAFU curriculum, instruction was designed to support students in moving from: identifying, describing, and symbolizing all of the symmetries of an equilateral triangle (task setting activity), to developing a calculus for combining symmetries (referential activity), to using the algebraic structure of this particular group of symmetries to analyze other systems and defining *a group* (generalized activity), to leveraging the properties of groups in order to develop the isomorphism concept (formal activity). Using this hypothesized progression as a guide, analysis can be carried out to determine the extent to which the students' activity followed this path. Tracing such a development would be evidence of changes in the students' participation and therefore would be evidence of student learning.

In addition to looking at changes in the students' mathematical practices by analyzing the trajectory of the student activity, it is also possible to look at changes in the mathematical practices by analyzing changes in the mathematical context in which the practices are taking place. The notion of a new mathematical reality (as understood from a participation perspective) provides a lens for describing the development of new mathematical contexts for further activity, and for describing new ways students participate in the new context. For instance, in the quotient group unit of the TAAFU curriculum, the guided reinvention design heuristic informed a progressive mathematizing sequence. This sequence culminates with an expanded mathematical reality that includes a working definition of quotient groups. Additionally, this sequence provides students with opportunities to reason within this new context. As students try to build partitions that form a group, they often try a number of different partitions and begin to develop a process for building quotient groups and an intuition about why some partitions form groups while others do not. In this way, the mathematical reality for the students' activity changes as they engage with the instructional sequence – both in terms of the context in which the activity takes place (an expanded context which includes quotient groups) and in terms of the way that the students interact in the context (in terms of the ways students reason about partitions). Again, documenting such shifts provides evidence of student learning.

Therefore, when analyzing student learning from a participation perspective, the RME design heuristics provide powerful lenses for documenting student practice and changes in these practices. The various mathematizing activities described in the literature provide examples and characterizations of mathematical practices. Documenting student participation in such practices is a necessary component to documenting student learning. However, it is also necessary to understand changes in the students' practice. The RME design heuristics provide two avenues for analyzing changes in practice. Learning trajectories based on supporting students in progressive mathematizing and/or progressing through layers of generality provide a framework for analyzing how the mathematical practices of the students are changing in regards to the generality of their activity. Additionally, the notion of a new mathematical reality provides a way to discuss both changes in the context of the students' activity and changes in how students participate in this new context.

Evidence for Student Learning from an Acquisition Perspective

The documentation of student learning from an acquisition perspective focuses on the development of the mathematical concepts. With both the emergent models and guided reinvention design heuristics, the mathematical concepts develop as aspects of the students' mathematical activity become reified. Instead of considering the reification of a global concept, here I will consider a smaller grain size of analysis by discussing the documentation of local evidence of student learning. This approach is similar to the one taken by Rasmussen and Marrongelle (2006), who pointed out that, "connecting the model-of/model-for transition to reification is a strong requirement that typically accompanies extended periods of time" (p. 391). Therefore, Rasmussen and Marrongelle chose to analyze teaching practices on the day-to-day level by focusing on a version of the emergent model heuristic that did not require reification (*transformational records*). Similarly, when considering the emergent model heuristic, I will consider evidence of student learning on a local level. These local changes can either be 1) related to the form of the model the, as described by the chains of signification construct, or 2) related to the function of the model, as described by the record-of/tool-for construct. In the case of the guided reinvention heuristic, the goal is to find evidence of incremental expansions in what is experientially real for the students. From an acquisition perspective, this is understood as a creation of a new mathematical reality, where new mathematical objects become incorporated into the students' experiential reality. Documenting changes to the mathematical reality will focus on changes in the objects that arise as artifacts of progressive mathematizing.

The TAAFU curriculum launches in the context of the symmetries of an equilateral triangle. As seen in Figure 2, an early sign that emerges in this context is composed of a signifier (an initial set of inscriptions for the six symmetries) and a signified (the students' activity of physically manipulating an equilateral triangle). The curriculum then prompts students to generate a new set of symbols in terms of F and R, and in doing so supports the progression of the chain of signification. This new set of symbols represents a signifier in the next step in the chain of signification, with the earlier sign sliding under to become the object that is being signified by these symbols. This shift in the form of the model to one that is more powerful can be seen as a local change that is part of (and supports) the more global transition to a model for. Therefore, one sign sliding under a subsequent sign supports the reification of the global model by supporting shifts in the form that the model takes. As a result, documenting instances in which signs slide under is a way to capture local shifts in students' concept development and can provide evidence of student learning.

Once the students develop a common set of symbols using F and R, they are asked to determine the result of the combination of any two symmetries. An operation table initially emerges as a record-of the students' activity. Later, as the students argue that the identity element of a group must be unique, students may draw on the operation table as a tool-for constructing a proof. This shift in the function of a representation of the model, from an inscription to a instrument, can be seen as a local change that is part of (and supports) the more global transition to a model-for. Therefore, an inscription shifting from a record-of to a tool-for supports the reification of the global model by supporting shifts in the function that the model serves. As a result, documenting local instances of such transitions is a way to capture local shifts in students' concept development and can provide evidence of student learning.

The guided reinvention design heuristic, and the process of progressive mathematizing, provides a lens to document incremental expansions in the student's mathematical reality. The quotient group unit is launched with the assumption that both the behavior of the even/odd integers and the operation table for the symmetries of a square are experientially real for the students. From here, the first symbolic artifact is a partition of the symmetries of square into two sets. However, it is not until the students engage in vertical mathematizing (by proving that this partition forms a group) that this partitioning activity becomes a new type of object that is accessible to the students on an intuitive level (i.e., a special type of group with two subsets as elements). This expansion in the students' mathematical reality, which can be understood in terms of new mathematical objects being accessible to students on an intuitive level, represents student learning from an acquisition perspective. Therefore, one way to document learning in this context is to look for evidence that new mathematical objects have become accessible and useful to the students as they work in more general problem contexts. In the example provided here, the students' activity of forming this new type of group (with two subsets as elements) resulted in a new object within the students' experiential reality (where the new object is the new type of group). In order to document such a change in the students' mathematical reality, one could look for evidence that this new object has become available for further progressive mathematizing. This could include students being able to further mathematize this expanded context to move beyond a focus on parity by intentionally forming groups made up of subsets.

So, when analyzing student learning from an acquisition perspective in situations where the learning is designed to be supported through a model of/for transition, we can look for local shifts in the form and function of the emerging model. This includes looking for indications that one sign has slid under a subsequent sign and looking for indications that a record-of student activity is serving as a tool-for subsequent student activity. Both of these local shifts support the reification of the global model (i.e., student learning from an acquisition perspective). Further, when analyzing student learning from an acquisition perspective in situations where the learning is designed to be supported through progressive mathematizing, we can look for incremental additions to the students' mathematical reality. These additions reflect that aspects of the students' activity have become objects that are now accessible for further mathematizing.

Conclusions

RME offers curriculum developers with a powerful theory for instructional design. Emergent models supports instructional design efforts by describing a mechanism through which students' informal and intuitive activity can be leverage to support the development of formal mathematics. Guided reinvention provides a description of how, by engaging in mathematical activity, students can expand the mathematical reality that they are able to access on an intuitive level. Part of the power of these two design heuristics resides in the fact that they place a dual emphasis on supporting both the students' mathematical activity and the formal mathematics that the curriculum is intended to develop. As a result, the curriculum developer can design curriculum with both student activity and concept development in mind.

While this focus on both activity and the concept development makes the theory flexible and powerful as an instructional design theory, it can be a confounding factor when trying to carefully articulate some of the fundamental RME constructs. For example, the idea of a *new mathematical reality* is left undefined although it is used often in the RME literature in order to describe the result of student mathematical activity. This lack of precision in the descriptions of these theoretical constructs became a significant problem as I tried to investigate the impact of the TAAFU curriculum on student learning. The curriculum was designed to support students in creating a new mathematical reality and developing formal mathematics through a model-of/model-for transition. So, it made sense to rely on these constructs to support my investigation. However, without knowing precisely what a *new mathematical reality* is, it is very difficult to argue that one has been established. In an effort to address such difficulties, this paper was written to explore the implications of RME for documenting student learning. I set out to first coordinate the RME theory related to the emergent model and guided reinvention design heuristics. Both of these heuristics support the development of new mathematical realities by engaging students in increasingly generalized activity, and both can be described in terms of more generalized activity and in terms of concept development. By focusing independently on these two aspects of the design heuristics, I was able to draw on Sfard's (1998) participation and acquisition metaphors for learning in order to discuss how these design heuristics support student learning.

Considering the design heuristics in light of these two perspectives on learning afforded a powerful lens for making sense of the idea of a new mathematical reality and for discussing what could be considered as evidence for student learning. I propose that, from a participation perspective, the creation of a new mathematical reality can be understood as the creation of a new context for further activity, and new ways for students to participate in that context. The RME design heuristics suggest a number of ways to document student learning from this perspective, including the following: documenting the mathematizing activities that students are engaged in; documenting how the mathematical practices of the students are changing in terms of the generality of their activity; and documenting changes in the students mathematical reality – both in terms of the context of the students' activity and in terms of how students participate in this new context. From an acquisition perspective on learning, I propose that the creation of a new mathematical reality can be conceptualized as the incorporation of new mathematical objects into the students' experiential reality. The incorporation of these new objects reflects that they have become accessible to students on an intuitive level. Again, the RME design heuristics suggest a number of ways to document student learning from this perspective, including the following: documenting when one sign has slid under a subsequent sign; documenting when a record-of student activity is serving as a tool-for subsequent student activity; and documenting incremental additions to the students' mathematical reality.

References

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (Vol. Lawrence Erlbaum Associates, pp. 307-334). Mahwah, NJ.

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3*(3/4), 413-435.

Freudenthal, H. (1991). *Revisiting Mathematics Education: China Lectures* Dordrecht, Netherlands: Kluwer.

Gravemeijer, K. (1999). Emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1*(2), 155-177.

Gravemeijer, K., & Doorman, M. (1999). Context problems in Realistic Mathematics Education: A calculus course as an example. *Educational Studies in*

Mathematics, 39, 111-129.
Gravemeijer, K., & Terwel, J. (2000). Hans Freudenthal: A mathematician on didactics and curriculum theory. *Journal of Curriculum Studies 32*(6), 777-796.
Johnson, E. (2013). Teacher's mathematical activity in inquiry-oriented instruction. *Special Issue of the Journal of Mathematical Behavior.*
Johnson, E. M. S., & Larsen, S. (2012). Teacher listening: The role of knowledge of content and students. *Journal of Mathematical Behavior, 31*, 117 - 129.
Larsen, S. (2004). *Supporting the Guided Reinvention of the Concepts of Group and Isomorphism: A Developmental Research Project* Unpublished Dissertation Arizona State University
Larsen, S. (2009). Reinventing the concepts of groups and isomorphisms: The case of Jessica and Sandra *Journal of Mathematical Behavior, 28*(2-3), 119-137.
Larsen, S. (2013) A local instructional theory for the guided reinvention of the group and isomorphism concepts. *Special Issue of the Journal of Mathematical Behavior.*
Larsen, S., Johnson, E., & Weber, K (Eds). (2013). The Teaching Abstract Algebra for Understanding Project: Designing and Scaling up a Curriculum Innovation. [Special *issue]. Journal of Mathematical Behavior.*
Larsen, S. & Lockwood, E. (2013) A local instructional theory for the guided reinvention of the quotient group concept. *Special Issue of the Journal of Mathematical Behavior.*
Rasmussen, C., & Blumenfeld, H. (2007). Reinventing solutions to systems of linear differential equations: A case of emergent models involving analytic expressions. *Journal of Mathematical Behavior, 26*, 195-210.
Rasmussen, C., & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics in instruction. *Journal for Research in Mathematics Education, 37*(5), 388-420.
Rasmussen, C., Zandieh, M., King, K., & Teppo, A. (2005). Advancing mathematical activity: A practice-oriented view of advanced mathematical thinking. *Mathematical Thinking and Learning 7*(1), 51-73.
Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.
Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4-13.
Wawro, M., Sweeney, G., & Rabin, J. M. (2011). Subspace in linear algebra: Investigating students' concept images and interactions with the formal definition. *Educational Studies in Mathematics, 78*(1), 1-19.
Weber, K., & Larsen, S. (2008). Teaching and Learning Group Theory. In M. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* MAA.
Zandieh, M., & Rasmussen, C. (2010). Defining as a mathematical activity: A framework for characterizing progress from informal to more formal ways of reasoning. *Journal of Mathematical Behavior, 29, 57-75*

WHY ADVANCED MATHEMATICS LECTURES OFTEN FAIL

Kristen Lew[1], Tim Fukawa-Connelly[2], Juan Pablo Mejia-Ramos[1], and Keith Weber[1]
Rutgers University[1] and Drexel University[2]

Research on mathematicians' pedagogical practice in advanced mathematics is sparse. The current paper contributes to this literature by reporting a case study on a mathematics professor's presentation of a proof in a real analysis course. By interviewing the professor, we focus on his pedagogical goals when presenting this proof and link those with the actions that he took to achieve these goals. By interviewing six students, we investigate how they interpreted the proof and what they learned from it. Our analysis provides insight into why students did not learn what the professor intended to convey in his presentation.

Key words: Lecture; Proof; Undergraduate mathematics education; analysis

Introduction and Research Questions

In advanced mathematics courses, what do mathematics professors try to convey in their lectures? How do they try to convey this mathematics to their students? How do students interpret the lectures? If students do not learn the mathematical content that the professors aim to communicate, what can account for their failure? Given the widespread prevalence of lectures in advanced mathematics courses, these questions are of central importance to research in undergraduate mathematics education. Yet, surprisingly, there has been relatively little empirical research in this area. The broad aim of this paper is to investigate these issues in the context of one mathematics lecture. To investigate this issue, we studied the presentation of a proof in an advanced mathematics class from three perspectives: the researcher's perspective, the professor's prospective, and the students' perspective. We used this to address the following questions:

(1) What mathematical content was the professor trying to convey in his presentation of the proof and how did he try to convey this content?

(2) How would knowledgeable members of the mathematical community interpret this proof? Would their interpretations align with the professor's?

(3) How did students who attended the lecture interpret this proof? Did their interpretations align with those of the professor or the mathematical community?

(4) If students did not interpret the lecture in the way the professor intended, what factors might have contributed to this unsuccessful communication?

Because this is a case study, this report will not contain findings that are necessarily generalizable. However, since these questions have rarely been addressed empirically in the literature (especially (3) and (4)), this report will offer important hypotheses about students' and professors' perceptions of mathematics lectures.

Related Literature

Perceptions of Lectures in Advanced Mathematics

Although lectures are the dominant form of pedagogy in advanced mathematics classrooms, many mathematics educators and some mathematicians question their effectiveness (e.g., Davis & Hersh, 1981; Dreyfus, 1990; Leron & Dubinsky, 1995; Hersh, 1993; Thurston, 1994). There are three common complaints about lectures in advanced mathematics classrooms. First, lectures emphasize formalism, sometimes to the point of consisting entirely of definitions, theorems, and proof. This emphasis denies students the opportunity to see informal modes of mathematical reasoning, such as how mathematical concepts might be represented diagrammatically, how proofs were generated, or why concepts are defined the way they are. As a result, students end up with a misleading view about mathematics and are unable to complete some important mathematical tasks such as exploring, defining and conjecturing (e.g., Davis & Hersh, 1981; Dreyfus, 1990; Thurston, 1994). Second, formal proof is a poor way to convey mathematical content and other less formal explanations might be more accessible to students (Leron, 1983; Hersh, 1993; Rowland, 2001). Third, mathematicians are disinterested in their teaching and hence do not put adequate effort into preparing their lectures (e.g., Davis & Hersh, 1981), perhaps because they believe their students are incapable of learning the material (Leron & Dubinsky, 1995). It should be emphasized here that these views are based on widespread opinion or some authors' own experiences rather than empirical studies of mathematicians' teaching or their perceptions of teaching.

What is Known About Lectures in Advanced Mathematics

In 2010, Speer, Smith, and Horvath claimed that there was only one empirical study in the literature that systematically examined what occurred in lectures in advanced mathematics while also interviewing the professor about his intentions (Weber, 2004). Since their review, several other mathematics educators have observed and analyzed mathematics lectures (e.g., Fukawa-Connelly, 2012, in press; Fukawa-Connelly & Newton, in press; Mills, 2014), but the instructor's and students' perspectives of these lectures was not investigated. Rather the authors imputed motives to the instructors to make sense of their behaviors and broadly described learning opportunities that students would have when observing the lectures. In general, the results of these studies, as well as the study of Weber (2004), are inconsistent with the widespread complaints against lectures. The lectures that were observed were not purely formal; professors regularly used diagrams (Mills, 2014; Weber, 2004) and examples (Fukawa-Connelly & Newton, 2012) and tried to model appropriate mathematical behaviors (Fukawa-Connelly, 2012). The professor interviewed by Weber (2004) cared deeply about his instruction and his lectures were based on a good deal of thought. It must be emphasized that these case studies were based on a small and possibly unrepresentative sample of mathematicians; professors who agree to have their lectures observed might have a tendency to be especially thoughtful and inclusive of informal explanations and examples. Nonetheless, these studies indicate that more research is needed on whether and why lectures are ineffective for students.

Theoretical Perspective

In this paper, we draw on three theories. First, we adopt several perspectives from the New Literacies Movement (Gee, 1990). We do not think of a lecture as being equivalent to a transcript of what was spoken, but rather treat the totality of a lecture, including the words spoken by the professor, the intonations of these words, chalk inscriptions, and kinesthetic movements, as a single coherent piece of text. We also follow Selden and Selden (2003) in adapting the framework of Bogdan and Straw (1990) to characterize three locations where the meaning of a mathematical text might reside: with the author of the text (i.e., the professor), with the reader (i.e., the student), or independent of the author and reader. It is interesting to note that mathematicians tend to treat mathematical text, and especially mathematical proof, as being independent of both the author and the reader. As Selden and Selden (2003) remarked, "mathematicians say that an argument proves a theorem, not that it proves it for Smith and possibly not for Jones" (p. 11) and Shanahan, Shanahan, and Misischia (2013) wrote that mathematicians actively try to avoid considering who wrote a mathematical paper when they read it[1]. As we noted previously, previous researchers on lectures in advanced mathematics implicitly adopted this position when they study mathematics lectures. They study what actually occurred during lecture, but do not actively consider the perspectives of the lecturer or student. In this paper, we do not suppose that there is a correct or best way to characterize where the meaning of a text resides. Rather we analyze the meaning of the text from three perspectives—that from the author of the text (i.e., the professor who presented the proof), the intended readers of the text (i.e., the students who observed the presentation of the proof), and from the perspective of knowledgeable members of the mathematical community.

Second, we follow de Villiers (1990) and others who observe that conviction is not the only reason, or even the primary reason, that proofs are presented in mathematics or university classrooms (e.g., Hanna, 1990; Hersh, 1993; Rav, 1999; Yopp, 2011; Weber, 2012). de Villiers (1990) listed four other functions of proof: explanation, discovery, communication, and systematization. For *explanation*, Hanna (1990) and Hersh (1993) argued that explanation should be the primary function of proof in the classroom. Weber (2010) suggested that for pedagogical purposes, one could view a proof as explanatory if students are able to relate the content of the proof to informal representations of mathematical concepts (such as diagrams, graphs, or kinesthetic motions) that are internally meaningful to the audience of the proof. For *discovery*, proofs of non-routine theorems often introduce new ideas or methods that can be extrapolated to discover and prove other theorems. Indeed, mathematicians claim that the primary reason that they read published proofs is to identify methods that will help them solve problems that they are working on (e.g., Rav, 1999). Hanna and Barbreau (2008) argued that classroom proofs could be more pedagogically valuable for students if they introduced new proving or problem-solving methods as well. By *communication*, de Villiers (1990) meant that by using proofs, mathematicians adopted shared standards of communication that could facilitate debate and resolution, but a discussion of these standards were not present in the lecture we observed or the interview with the professor. Likewise, *systematization*, or axiomatizing a theory, was not observed.

Third, we adopt the theoretical notion of codes proposed by Weinberg and Weisner (2011) on the reading of mathematical text. The meaning of a text is not literally contained in the text. Rather, the author of the text encodes his or her intended meaning in some way. We refer to this as the encoded content of the text (throughout the paper, we define "content" broadly to be anything that the author of a proof intended to convey or anything the reader of the proof took away from it). Weinberg and Weisner (2011) argued that readers of texts have codes, defined to be the reader's "system of signification" or "a way of ascribing meanings to the parts of the text" (p. 53). If a reader of a proof had different codes than the intended universal audience of the proof, his or her interpretation of the proof will differ from that of the author's and hence will not develop the mathematical understandings and insights that the author intended to convey. Of course, possessing the appropriate codes does not guarantee that comprehension will occur, but lacking them will make comprehension difficult.

Related Literature

The Lecture

This research took place at a large state university in the northeast United States in a real analysis course. At this university (and most universities in the United States), real analysis is a junior-level course that is required for mathematics majors. We chose to study a section of

[1] Based on empirical studies, we believe these claims and those that follow are overstated (Inglis et al, 2013; Weber, Inglis, & Mejia-Ramos, 2014; Weber & Mejia-Ramos, 2011).

the course taught by Dr. A. Dr. A had extensive experience teaching collegiate mathematics and regularly taught this real analysis course. He also received very high teaching evaluations from his students and had a reputation for being an outstanding instructor amongst his colleagues.

The lecture that will be investigated in this paper was videotaped during the eighth week in the course. During the lecture, Dr. A almost always positioned himself between the students, who were sitting in desks, and the blackboard. Our video recording of the lecture focused exclusively on the actions of the professor, capturing what comments he said orally, what he wrote on the blackboard, and the gestures that he made. Our analysis in this paper focuses on a proof about sequences $\{x_n\}$ having the property that there exists a constant r (with $0 < r < 1$) such that for any two consecutive terms in the sequence x_{n-1} and x_n, it is the case that $|x_n - x_{n-1}| < r^n$. Dr. A proved a theorem that sequences with this property are convergent. This 11-minute proof presentation was one of seven proofs in the lecture and was chosen because it was the most conceptually interesting; the other six proofs focused on computation or showing that a particular example satisfied a given definition.

To avoid ambiguity, we refer to the *blackboard proof* as the argument that Dr. A inscribed on the blackboard demonstrating that the theorem was true. We refer to the *lecture proof* as the totality of the prof, including oral comments and gestures made by Dr. A. The focus of this study is on the lecture proof.

Our Interpretation of the Text

As a first pass through the data, we attempted to interpret what content could be learned through the lecture from by a mathematically enculturated individual. The lecture proof was viewed individually by the four authors of the paper, who all had experience teaching courses in advanced mathematics and all had or were pursuing master's degrees in mathematics.

Each member of the research team flagged instances when he or she felt that Dr. A was trying to convey an idea to his students. For each instance, the researcher noted what content was being covered and how it was encoded. We also coded the type of content based on de Villiers' (1990) purposes of proof. If the emphasis of the content was on verifying that a given statement was true, we coded this content as an instance of *verification*. If Dr. A gave a conceptual explanation for why the theorem was true, we coded this content as *conceptual explanation*. If Dr. A highlighted ideas within the proof that might be useful for discovering or proving other theorems, we coded this content as *method*. After individually coding the lecture, the research team met to compare their findings and reach a consensus. We sought independent confirmation of our analysis by asking a lecturer in mathematics who was currently teaching a course in real analysis to view the videotape and describe what he thought were the main ideas of the lecture.

The Author's Aims and Interpretation of the Text

After the initial analysis of the text, the first author met individually with Dr. A for an audio-recorded interview. The interviewer first asked Dr. A why he chose to present this theorem and its proof. Dr. A was then asked what he thought were the main ideas he was trying to convey to the class in his proof. Next, Dr. A was shown the video from his lecture of the proof and was asked to stop the recording at any point where he was attempting to convey the content that he just described, or to identify any other points that he may have neglected to mention. Whenever Dr. A stopped the tape to identify content that he was trying to communicate to students, the interviewer would ask how Dr. A was trying to convey that content to the students. This interview lasted 75 minutes.

We analyzed Dr. A's comments about the content he was trying to convey using a semi-open coding scheme. If his comments were consistent with what we observed, we would fold them into the categories that we formed in our analysis of the tape. If not—that is, if he introduced new content or described the content that we observed in a different way—we would form a new category. Again, the content of this category was then coded using the categories that we described above. We also noted where, if at all, these ideas were encoded in the proof he presented, both by studying the points where he chose to stop the video and through viewing the proof in its entirety on our own.

The Students' Interpretation of the Text

After the lecture was given, the first author went to Dr. A's course and invited students to participate in a study on how they understood a mathematical lecture. Students were paid a nominal fee their participation. Six students volunteered to participate in this study. Two weeks had elapsed between the interviews and the lecture. The reason for this delay is that Dr. A forbade the research team from recruiting subjects prior to that point as the students were being given their mid-term examination.

The students were interviewed in pairs, as we anticipated that the opportunity to communicate with one another would elicit more comments from the students. We refer to the first pair of interviewed students as Pair 1 and the individual students as S1 (Student 1) and S2 and so on. Each pair of students was asked to bring their lecture notes to the interview, which the interviewer photocopied. All interviews were video-recorded. The interview involved four passes to explore students' understanding of the proof. The data collection and intention of each pass through the data is presented in Table 1.

Table 1.
Summary of four passes through the lecture proof with students

Pass	Data collection	Purpose
Pass 1	Participants recalled what they learned from the proof based on their notes.	We wanted to see what participants could reconstruct from a lecture proof after time had passed.
Pass 2	Participants viewed the video recorded presentation of the proof, took notes, and were asked what they learned and what the instructor was attempting to convey.	We wanted to see what participants understood immediately after viewing the proof.
Pass 3	Participants watched short specific clips from the proof and were asked what Dr. A was trying to convey.	We were investigating whether participants had the specific codes needed to interpret what Dr. A had identified as the content in his presentation of the proof.
Pass 4	Participants were asked whether particular content highlighted by Dr. A in his interview could be gleaned from the proof they just watched.	We were exploring how participants understood the main ideas that Dr. A claimed he was trying to convey with this proof.

In the first pass, students were invited to review their notes, and asked to describe what they thought were the main ideas of the lecture. In the one occurrence that a student (S2) had not taken notes during lecture he was provided with a copy of everything Dr. A wrote on the board, which was distributed to the rest of the students in the second pass. This first pass through the data was designed to measure what students could reasonably reconstruct from their experience attending a lecture. As a second pass through the data, students were asked to review the videotape of the proof in the lecture in its entirety. They were asked to behave as if they were in a mathematical lecture, including taking notes. However, they were also given a copy of everything that was written on the board, as the blackboard proof could be difficult to read while watching the video recording. After watching the video, they were then asked to describe what they thought the professor was trying to convey when he was presenting the proof. The second pass through the data was to see what students understood about the proof immediately after watching it.

For the third pass through the data, the participants were shown each of the particular clips that Dr. A flagged as points where he was trying to convey some specific content. The participants were told these excerpts were places where Dr. A thought he emphasized some important ideas in the proof. The participants were asked to describe what they thought about these clips, but the interviewer added, "it's acceptable to say that you don't see anything. I don't want to encourage guessing". The goal of this pass through the data was to see how they interpreted specific episodes of the lecture, and to determine to what extent the participants' interpretations matched what Dr. A was trying to convey in these episodes. This was an attempt to measure the extent to which the participants had the codes needed to interpret what Dr. A considered to be the main content of the proof. In total, the participants saw eight excerpts.

In the final pass through the data, the interviewer identified each type of content that Dr. A claimed he was trying to convey in this proof. The participants were told that some versions of this proof could convey this specific piece of content and asked if they thought the proof that they just observed did this. For instance, in his interview, Dr. A stressed that it was useful to think of the terms of the sequence as approximations of the limit and the epsilon in the definition of the limit as representing the error of the approximation. The question the interviewer asked with respect to this content was, "another thing you might have gotten from this proof is the epsilon used is the error. Is that something that you got from this presentation?" If the participants answered affirmatively, the interviewer encouraged them to describe how the proof conveyed this content and how they understood the content. The goal of this pass through the data was to see if, and to what extent, participants could access and understand the main points that Dr. A was trying to convey if asked specifically about them.

In analyzing the first and second passes through the data, we used open coding to determine what these three pairs of students identified as the important ideas in the proof; our initial categories of the content were those formulated by our research team and claimed by Dr. A, but we formed new categories if the students' comments did not fit within our initial framework. In the third pass through the data, we compared students' interpretation of the video clips to the meaning that Dr. A ascribed to them in his interview. In the fourth pass through the data, we analyzed if students' understanding of specific content of the proof was aligned with Dr. A's intentions.

Results: Our Analysis of the Lecture

General Comments

This lecture proof had characteristics that were typical of many lectures in advanced mathematics. Dr. A spent his entire time between the students and the blackboard; for the majority (61%) of the time, his back was to the class as he was writing the proof on his blackboard. In our judgment, students only had minimal participation in the proof. At five points during the proof, Dr. A asked questions of the class, but two of those questions seemed rhetorical since Dr. A quickly provided the answer to them before students had the chance to respond. The other three questions used an initiate-response-evaluate format (Mehan, 1979); in each case, Dr. A asked students to supply the next step in the proof that was being worked on. Twice the students did not immediately supply an answer to Dr. A's question so Dr. A provided a hint. For instance, the following exchange occurred:

> Dr. A: Now we know this is small [circling one mathematical expression]. Now what can we say about this expression right here [circling another mathematical expression]? [pause] Anybody have a vague idea? I'll give you a hint. Calculus two...
> Student: Geometric series?
> Dr. A: … thirty or forty years ago? [gestures to student who spoke]
> Student: Geometric series.
> Dr. A: Geometric series! You have to always keep geometric series in your toolbox.

We also noted that Dr. A's lecture proof was significantly more detailed than his blackboard proof. The blackboard proof consisted of a polished proof that might appear in a textbook. However, in the lecture proof, he supplemented this with oral comments describing methodological and conceptual explanatory content.

Main Themes from the Lecture

In the lecture, we noted four main ideas that Dr. A was trying to convey:

(i) One can show a sequence is convergent by showing it is a Cauchy sequence, which is especially useful when you do not know what the limit of the sequence will be.

(ii) There is a common structure for writing proofs showing sequences are Cauchy.

(iii) The triangle inequality is useful for showing that the sum of small terms is small.

(iv) The geometric series formula is a useful technique for working with inequalities in real analysis. It should be in a student's mathematical toolbox for keeping quantities small.

We coded each one of these ideas as methodological content as all were useful in helping students write proofs in the future. Due to space limitations, we will only discuss (i) and (iv). Recall the theorem being proven was that if a sequence $\{x_n\}$ had the property that $|x_n - x_{n-1}|<r^n$ for some constant r between 0 and 1, then $\{x_n\}$ was convergent. Previous proofs establishing convergence typically were based on the definition of limit, but using this definition required having a limit candidate for the sequence, something that could not be done in this situation because of the arbitrariness of the sequence. To us, (i) was the key theme from this part of the lecture: One could establish convergence by showing a sequence was Cauchy, even if one did not have a candidate for the limit. Dr. A emphasized this at three different points in the lecture, as we illustrate in the two excerpts below:

> Dr. A: There's no mention of what the definition is of the sequence, so there's no way we're going to be able to verify the definition limit of a convergent sequence, where we have to produce the limit. So what do we do? […] What kind of sequences do we know converge even if we don't know what their limits are? It begins with a 'c'.

> Dr. A: This is how we prove it is a Cauchy sequence. See there is no mention of how the terms of the sequence are defined. There is no way in which we would be able to propose a limit L. So we have no way of proceeding except for showing that it is a Cauchy sequence.

Regarding (iv), after Dr. A invoked the geometric series formula in his proof, he said, "geometric series! You have to always keep a geometric series in your toolbox". As individuals who have completed and instructed a real analysis course, we understood the phrase "mathematical toolbox" in this setting to mean a collection of techniques for working with inequalities to keep desired quantities small. We believe this shared understanding is common among mathematically enculturated individuals in this context. For instance, in a separate study, when Weber (2004) asked another professor about his goals when presenting this type of proof in real analysis, the professor said, “I would like my students to have a *mathematical toolbox*. If they know the limit of something exists, they should immediately think of ways to make the desired quantities small” (p. 124, italics were from the original paper). We corroborate that this was Dr. A’s interpretation of “toolbox” in the next section. Of course, this interpretation likely would not be apparent to someone who was not experienced in this subject, including students in a real analysis course.

Comments from Another Course Instructor

To corroborate our findings, we showed another course instructor the video recording of the lecture. The instructor was extremely complimentary of the quality of the lecture and thought the main ideas of the lecture were (i), (ii), and (iv). He did not mention (iii), the importance of the triangle inequality and he did not mention other content that we did not. Like us, he thought (i) was the main goal of the proof, saying, “I think that’s the main objective here … up until this point, they’ve been showing convergence by definition… So he's saying okay, but what if you don't have a way to find the limit? Can you still show something converges? Well yeah, if you know it's Cauchy. I think that's one thing he's going to be trying to do here”.

Key Points from this Section

Dr. A conveyed important methodological content in his proof. The lecture did not seem geared toward convincing students that a theorem was correct or proving to satisfy a ritual, but toward helping students know how to construct proofs in the future. This is consistent with the teacher observations of Fukawa-Connelly (2012) and Weber (2004). We also note that the main content of the proof was stated orally, but not written on the blackboard.

Results: Dr. A’s Analysis of the Lecture

General Comments

The interview with Dr. A was surprisingly long. He spent 75 minutes discussing a proof that took him ten minutes to present. There are several things noteworthy about this interview. First, when asked why he chose to present this proof, Dr. A gave an 11 minute account situating Cauchy sequences in students’ analysis learning trajectory starting with calculus up through the study of measure theory in graduate school, suggesting that he had thought carefully about the place of this proof in instruction. Second, he emphasized the importance of repeating themes over the course of the semester, both to give students multiple opportunities to learn the same material and to form automatic associations between mathematical concepts, mental imagery, and proving actions.

Methodological Content

When Dr. A was viewing the videotape, he stopped the tape at every point that we identified mathematical content and his description agreed with our interpretation. One particular excerpt concerned his description of the triangle inequality.

> Dr. A: Once you get into the area where you're doing approximations, you can't do equal, equal, equal. You have to have bounds, bounds, bounds […] *The objective is to show how bounds, using the triangle inequality, can be used to show that something is small using information that they're given is small.* And this instance turns out that the information which is small is given in a form that allows us to use the geometric series as a bound. (italics were our emphasis).

We highlight this as being consistent with our interpretation that it is important for students to have techniques to find bounds that keep the sums of small quantities small.

Conceptual Explanatory Content

Dr. A discussed several types of content that we coded as conceptual explanatory, as he sought to explain the theorem and illustrate the concepts using visual, kinesthetic, or metaphorical reasoning. In particular, he said a goal of this proof was to illustrate Cauchy sequences pictorially, to represent Cauchy sequences as "bunching up", to think of the epsilon term as an error of an approximation, and to view proofs about convergent sequences as analogous to knowing when a computer algorithm that forms approximations should terminate.

Despite being listed as content he wished to convey in this proof, this content was nearly entirely absent from the proof. Only the notion of Cauchy sequences "bunching up" was evident with a brief gesture showing his hands coming together (we missed this in our analysis, but Dr. A stopped the tape to highlight where this occurred). It is notable that no picture was given in the proof. In our interview, Dr. A spent six minutes describing the importance of pictures in real analysis, using the word "picture" 32 times, and saying students needed to see pictures at every opportunity so they would automatically associate Cauchy sequences with a picture. But when shown the tape, Dr. A laughed and said, "this is a poor example. There are no pictures here!"

Key Points from this Section

Dr. A appeared to (or at least talked as if he) thought hard about his lectures and was concerned about his students' development. Dr. A valued repetition in his lectures so that students would form associations between concepts and pictures as well as to give students multiple opportunities to learn the material. This finding was also present in the literature by Weber (2004, 2012) in a case study of one mathematician's teaching and interviews with mathematicians. Dr. A reported having the conceptual explanatory goals valued by mathematics educators, but this content was notably absent from his lecture proof. This is consistent with findings from Alcock (2010) and Lai and Weber (2014), who noted that mathematicians do not include all the conceptual content that they plan or desire to include in their lectures.

Results: Students' Analysis of the Lecture

Pass 1

We photocopied students' notes from the class. One student had a near verbatim description of everything that Dr. A said, four students copied the blackboard proof but nothing else, and one student did not take notes that day (but did attend the lecture). We find it noteworthy that five of the six students did not transcribe the oral content of the proof that Dr. A thought was so critical for conveying the mathematical content of the proof.

In this pass, students were asked to look over the proof in their notes and asked what they thought Dr. A was trying to convey. Perhaps unsurprisingly, their responses at this stage were shallow. Pair 1 described a heuristic for knowing when to prove something was Cauchy-- namely if subscripts with the variables m and n were present-- and that Cauchy sequences can be used to prove convergence. One student in Pair 2 seems to have misinterpreted the proof, viewing the set-up for the proof and the algebraic manipulations in the proof as "two different methods to drive home the same thought process" for "people [who] learn differently". Pair 3 was excited to talk about the proof, with both students instantly recalling the proof used geometric series. Their description focused on the algebraic methods used to simplify equations, noting that different techniques were needed for different problems and one can rely on background knowledge from other courses to do this.

We note that no student mentioned the main point: Proving a sequence is convergent by proving it is Cauchy is particularly useful when one does not know what the limit of the sequence is. Although Pair 3 highlighted techniques for working with inequalities, no students spoke of techniques that used bounds to keep quantities small.

Pass 2

The students were asked the same questions as they were in Pass 1 after watching a video of the lecture proof. Students' performance improved; they highlighted aspects of the proof that were both correct and useful. Pair 1 highlighted that (a) Cauchy sequences will be on the mid-term (which Dr. A stated in the proof), (b) the use of geometric series brought in prior knowledge, and (c) the importance of using the triangle inequality. Pair 2 highlighted that (a) one can show a sequence is convergent by showing it is Cauchy, (b) there is a consistent structure to writing proofs about convergence, and (c) one can use ideas from calculus (specifically geometric series and the triangle inequality) to write proofs in analysis. Pair 3 highlighted that (a) the proof expanded the students' toolbox of how to simplify expressions and (b) the proof illustrated how students can use prior knowledge from calculus to write proofs in analysis.

There are a few interesting notes about these observations. First, all three groups noted the importance of using prior background knowledge in real analysis. Neither Dr. A nor our research team highlighted this as the mathematical content in the lecture but it is nonetheless a useful point. Again, no student said what we felt was the most important point of the lecture proof-- namely that if one wants to show a sequence is convergent but cannot determine its limit, one can do so by showing it is Cauchy.

Pass 3

In this pass, we showed students short video clips that Dr. A highlighted as conveying mathematical content. In general, students did fairly well in this pass. For instance, all accurately noted the importance of the triangle equality and saw Dr. A's hands clasping together as representing Cauchy sequences bunching up. For the sake of space, we discuss only two video clips.

In the first clip, Dr. A said, "What kind of sequences do we know converge even if we don't know what their limits are? It starts with a 'c'?" When a student replied, "Cauchy", Dr. A said, "Cauchy! We'll show it's a Cauchy sequence". Both Pair 1 and Pair 2 believed Dr. A was trying to convey that one can show a sequence is convergent by showing it is Cauchy, which is useful if you do not know the limit of the sequence. For instance, S1 said, "we should recognize it, like to figure out it's a Cauchy, we should know that it's converging, but its limit is not necessarily given. So that we recognize it instantly". However, Pair 3 thought the purpose of the clip was for Dr. A to interact with the class by posing a question and to gauge what the students knew. In fact, in Pass 3, we showed three video clips where Dr. A emphasized that showing a sequence is Cauchy could establish convergence if the limit was unknown and Pair 3 never thought that was what Dr. A was trying to convey.

In the second clip, Dr. A said the following:

> Dr. A: So let's factor out the smallest term, r to the n. What's left is 1 + r + r squared + up to r to the m - n. [Writes this equation on the blackboard as he speaks]. Now we know this is small [circles r^n] now what can we say about this expression right here? [points to and circles the geometric series $1 + r + r^2 + \ldots + r^{m-n}$, then turns around and faces the class]. Anybody have a vague idea? I'll give you a hint: Calculus II. Thirty or forty years ago.
>
> Student: Geometric series.
>
> Dr. A: Geometric series! [Turns and faces the blackboard]. You have to always keep a geometric series in your toolbox. So it's going to be less than r^n, this [gestures towards the geometric series written on the blackboard] then is less than sum from k=0 to infinity of r to the k. And now we need to know the formula of a sum of a geometric series.

When asked what Dr. A was trying to convey, no student mentioned the notion of toolbox. Rather, all three pairs of students highlighted the importance of referring back to previous knowledge to complete this proof. For instance, S6 stated Dr. A was "trying to like convert this expression somehow into an expression that we are familiar with or we know about from like our previous courses -- in this case it would be geometric series".

Pass 4

Dr. A listed a number of mathematical ideas that he was trying to convey. In Pass 4, for each purpose that Dr. A listed, each pair of students was asked, "one thing you might get from this proof is that [Dr. A's purpose]. Is this something that you got from this presentation?" Often, for these questions, students would respond "yes", but when elaborating, it became apparent that their meaning of the terms used in Dr. A's purpose differed from Dr. A intended.

For instance, students were asked if the proof could provide students with "a toolbox that help them to prove things are small". All students answered "yes" to this question. . However, in their responses, none mentioned inequalities or making things small. Indeed, from their responses, it appeared that students viewed the components of the toolbox as general techniques for writing proofs in mathematics. S2 described Cauchy sequences as being part of his toolbox, indicating the toolbox was for proving convergence, not keeping quantities small. S3 described, "I think if he structures the way that he does, and you keep seeing it, it stays in your toolbox memory area […] not just in this specific proof itself, but it carries over to any other areas of math when you want to start to prove something". Again, here it appears that how one structures real analysis proofs is part of one's toolbox.

Only one student, S5, mentioned the word "small" in his or response. In the following excerpt, we can see that S5 was not using small in terms of a magnitude of a quantity, as Dr. A intended.

> S5: We can use Mathematica, or like a tool to convert to make something small.
>
> I: So right so mathematics students need to have a toolbox of ideas to help them prove things are small.
>
> S5: Things are small. Oh you mean that they're not so complicated. When you say that things are small?
>
> I: No I mean like in terms of convergent sequences. Is that something that you think you got from this presentation?
>
> S5: I mean, in terms of simplifying them and deriving for approximating the answer, I think it's on the path, it's like it's working.

Again, we see that by listing Mathematica (a computer algebra system commonly used in college calculus classes but not real analysis), S5 is referring to general mathematical tools, rather than tools for working with inequalities or keeping quantities small. His response to the next question revealed that S5 did not know what was meant by "small" in this context, guessing that it means a not complicated, or simplified, equation, rather than a quantity with a small magnitude.

Similar responses were made by students when they were asked if "epsilon as error" was conveyed in this proof (all students agreed but none alluded to approximations in their responses) or applications to computer science were conveyed in this proof (five students agreed, but none explicated any analogy between the two domains).

Key points and summary.

In Table 2, we list each of the types of content that Dr. A believed he encoded in his lecture and the pass through the data that students first articulated this content.

Table 2.
Summary of when student pairs described Dr. A's intended content

Content	Pair 1	Pair 2	Pair 3

Cauchy sequences can be understood as sequences that "bunch up"	Pass 3	Pass 4	Pass 3
One can prove a sequence with an unknown limit is convergent by showing it is Cauchy	Pass 3	Pass 3	Never
How one sets up a proof that shows a sequence is Cauchy	Pass 4	Pass 2	Pass 4
The triangle inequality is useful in proving series in absolute value formulae are small	Pass 2	Pass 3	Pass 3
The geometric series formula is part of the mathematical toolbox to keep some desired quantities small	Never	Never	Never

Table 2 reveals several things. First, participants did not express many of the ideas in the lecture in their first two passes through the data. This is significant as the second pass through the data put the students in a privileged environment. They were watching a real analysis professor with an excellent teaching reputation present a proof for a second time. They knew they would be asked about this proof, so it was less likely that their attention would drift. If *these* students could not grasp the main points of the lecture by Pass 2, this suggests that *many* students are not grasping these main points with an average lecturer in real time. Second, participants did much better at Pass 3, implying that they had the codes to interpret much of Dr. A's oral comments. Still, the analysis of Pass 3 and Pass 4 suggests this was not always the case. For instance, Pair 3 never stated what we felt was main point of the proof despite seeing three clips where it was stated, and no students described the meaning of the toolbox in an accurate manner.

Discussion

Why Did Students Have Difficulty Understanding this Lecture in Advanced Mathematics?

We suggest several grounded hypotheses for students' difficulties:

H1. Dr. A did not include all the conceptual-explanatory content that he intended to, denying students the chance to see this material.

H2. The method content of the proof was stated orally but five students only transcribed the blackboard proof. Hence, the oral content valued by Dr. A was not recorded for further study by these students and may not have been attended to.

H3. There was some content, such as the notion of toolbox and small in this context, that students lacked the codes to interpret.

H4. Dr. A repeated several themes in the course. Students were able to recognize the words in this theme but not their meaning.

The point of raising these hypotheses is neither to blame Dr. A nor to blame the students for the state of affairs we observed. Indeed, all seemed well intentioned and we feel sympathy for both. For (H2), Dr. A might not have written down his oral comments since students at this stage were still learning what a proper proof should look like. Cluttering the blackboard with written translations of his oral comments might have distracted students from this worthwhile goal. Similarly, students may not have transcribed the oral comments since it may have been a struggle to simply transcribe the blackboard proof (one student mentioned this explicitly). For (H3) and (H4), it is not hard to imagine why students would think of one's toolbox as any useful mathematical technique rather than tools to work with small inequalities, as this meaning would make sense in other mathematical context (e.g. in a first or second calculus course, a toolbox might consist of integration techniques). Dr. A's repetition of themes is a natural pedagogical technique to emphasize importance and help students develop associations, but (H4) suggests this might not be enough. One plausible account might be that Dr. A is operating on the assumption that the rich mental images that he associates with the terms he uses may be immediately accessible to the students when, in fact, the students need to construct these mental images. If so, from the students' perspective, comprehending these terms might not be a matter of *association* but rather of *construction*. The latter process is likely too time-consuming and cognitively complex when hearing Dr. A's utterances in real time. Further, since students are not recording Dr. A's oral comments in their notes, they do not have the opportunity to construct these mental images by reflecting on Dr. A's comments for an extended period of time at a later point.

Significance for Research and Teaching

The obvious and important caveat to this study is that this was a single case study with a non-representative professor. Hence one certainly should not make general claims about lectures from these data. However, these data do make a useful contribution to the literature in two ways. First, this is the first study that we are aware of that studies lectures in advanced mathematics from both the professor's and the students' perspectives. Research on both perspectives is sorely needed to better understand the effectiveness of these lectures. Second, some hypotheses reported here lend themselves to systematic testing. For instance, it would not be difficult to see if mathematical content was usually stated orally, but not written down, by other professors, nor would it be difficult to collect students' notes and see how often they recorded oral comments. Also, although this would require more work, one could see if students' understandings of common metaphorical expressions in mathematics lectures, such as "epsilon is error" or "mathematical toolbox", are consistent with the professor who used these themes. If the grounded hypotheses from this study generalize to other lectures, this would provide valuable insights into why lectures are often ineffective and how this might be remediated. In summary, like many exploratory qualitative studies, the contribution here involve highlighting phenomena for future research and orienting researchers' attention to aspects of the phenomena that are likely to be important.

References

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in Collegiate Mathematics Education VII* (pp. 63-92). Providence, RI: American Mathematical Society.

Bogdan, D., & Straw, S. B. (Eds.) (1990). *Beyond communication: Reading comprehension and criticism*. Portsmouth, NH: Heinemann.

Davis, P. J. and Hersh, R. (1981). *The mathematical experience*. New York: Viking Penguin Inc.

de Villiers, M.D. (1990). The role and function of proof in mathematics. *Pythagoras, 24,* 17-24.

Dreyfus, T. (1990). Advanced mathematical thinking. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education.* Cambridge: Cambridge University Press.

Fukawa-Connelly, T. (2012). A case study of one instructor's lecture-based teaching of proof in abstract algebra. *Educational Studies in Mathematics*. DOI: 10.1007/s10649-012-9407-9

Fukawa-Connelly, T. (in press). Toulmin Analysis: A tool for analyzing teaching and predicting student performance in proof-based classes. *International Journal of Mathematical Education in Science and Technology.* DOI:10.1080/0020739X.2013.790509

Fukawa-Connelly, T., & Newton, C. (in press). Evaluating mathematical quality of instruction in advanced mathematics courses by examining the enacted example space. *Educational Studies in Mathematics.*

Gee, J. P. (1990). *Social linguistics and literacies: Ideology in discourses. Critical perspectives on literacy and education*. London: Falmer Press.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6-13.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM, 40*, 345-353.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, *24*(4), 389-399.

Inglis, M., Mejia-Ramos, J.P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science*, 5, 270-282.

Lai, Y. & Weber, K. (2014). Factors mathematicians profess to consider when presenting pedagogical proofs. *Educational Studies in Mathematics*, 85, 93-108..

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, *90*(3), 174-184.

Leron, U., & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly*, *102*, 227-242.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Mills, M. (2012, February). Investigating the Teaching Practices of Professors When Presenting Proofs: The Use of Examples. In *Proceedings of the 15th Conference for Research in Undergraduate Mathematics Education.* Edited by: Brown, S., Larsen, S., Marrongelle, K. and Oehrtman, M. (pp. 512-516). Portland, OR.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica, 7*, 5-41.

Rowland, T. (2001). Generic proofs in number theory. In S. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 157-184). Westport, CT: Ablex Publishing.

Selden A. & Selden J. (2003) Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34, 4-36.

Shanahan, C., Shanahan, T., & Misischia, C. (2011). Analysis of expert readers in three disciplines: History, mathematics, and chemistry. *Journal of Literacy Research*, *43*(4), 393-429.

Speer, N., Smith, J., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99-114.

Thurston, W.P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, 30, 161-177.

Weber, K. (2004). Traditional instruction in advanced mathematics classrooms: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, *23*, 115-133.

Weber, K. (2010). Proofs that develop insight. *For the Learning of Mathematics*, 30(1), 32-36.

Weber, K. (2012). Mathematicians' perspectives on their pedagogical practice with respect to proof. *International Journal of Mathematics Education in Science and Technology*, *43*(4), 463-475.

Weber, K., Inglis, M., & Mejía-Ramos, J.P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. Educational Psychologist, *49*(1), 36-58

Weber, K. & Mejía-Ramos, J.P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329-344.

Weinberg, A. & Wiesner, E. (2011). Understanding mathematics textbooks through reader-oriented theory. *Educational studies in mathematics 76*(1), 49-63.

Yopp, D. (2011). How some research mathematicians and statisticians use proof in undergraduate mathematics. *Journal of Mathematical Behavior*, *30*, 115-130.

EFFECTS OF SYSTEMATIC LISTING IN SOLVING COUNTING PROBLEMS

Elise Lockwood
Oregon State University

Bryan Gibson
University of Wisconsin-Madison

Although counting problems are easy to state and provide rich, accessible problem solving situations, there is evidence that students struggle with solving counting problems correctly. With combinatorics (and the study of counting problems) becoming increasingly prevalent in K-12 and undergraduate curricula, there is a need for researchers to identify potentially effective instructional interventions that might give students greater success as they solve counting problems. We tested one such intervention – having undergraduate students engage in systematic listing of what they were trying to count. We show that even creating partial lists of outcomes led to statistically significant improvements in students' performance on problems, implying that systematic listing may be worthwhile for students to engage in as they learn to count. Our findings suggest that instructional interventions that facilitate listing warrant more attention.

Key Words: Combinatorics, Systematic Listing, Counting Problems, Discrete Mathematics

Introduction and Motivation

The solving of counting problems has become increasingly prevalent in K-12 curricula (e.g., English, 2005) and in undergraduate mathematics courses. This attention on counting may be due to the fact that counting has practical applications in areas such as probability and computer science. Additionally, counting problems are simply stated and require few mathematical prerequisites to explore, and yet they require critical mathematical thinking to solve. This combination of accessibility and difficulty provides a uniquely rich context for mathematical problem solving (e.g., Kapur, 1970; Martin, 2001). In spite of the importance of counting problems, a number of studies that have been undertaken on combinatorics education suggest that students face difficulties with solving counting problems correctly. Given such struggles, there is a need for more investigations into effective ways to improve students' counting. In this paper, we share findings from a study that examined the effects of having students engage in *systematic listing* – that is, to create an organized list (or even a partial list) of the outcomes they are trying to count. We answer the following research questions:

1) Does engaging in systematic listing have a significant effect on students' solving counting problems correctly?
2) What are features of productive or unproductive lists that students generate, and how does listing activity differ in the generation of productive versus unproductive lists?

Literature Review and Theoretical Perspective

Students' difficulties with counting. While some researchers report success in which even young children display robust combinatorial thinking (e.g., English, 1991; Maher, Powell, & Uptegrove, 2011), most research on students' work on counting problems shows that students struggle substantially with solving counting problems. One piece of evidence of this difficulty is low overall success rates. Godino, Batanero, & Roa (2005) note that 118 undergraduate mathematics majors "generally found it difficult to solve the problems (each student only solved an average number of 6 [of 13] problems correctly)" (p. 4). Additionally, in their study on undergraduates' verification strategies, Eizenberg and Zaslavsky's findings "support the assertion that combinatorics is a complex topic – only 43 of the 108 initial solutions were correct" (2004, p. 31). In addition to low performance rates, there is evidence in the literature of how, specifically, students struggle. Batanero, Navarro-Pelayo, & Godino (1997) listed several error types they found in students' work. Eizenberg and Zaslavsky (2004) and Lockwood (in press a) point out that because of the nature of counting problems and their very large numerical answers, such problems can be difficult to verify. Other researchers have highlighted specific mathematical features of counting problems that are especially difficult, such as issues of order (Batanero, et al., 1997; Mellinger, 2004) and overcounting (Lockwood, 2011; Annin & Lai, 2010). Given the pervasive difficulties that students face, there is a need to identify potentially productive interventions that may help students solve counting problems more successfully.

Sets of outcomes. Theoretically, our focus on systematic listing stems from the idea that students may benefit from grounding their counting activity in the concrete *set of outcomes* they are trying to count. This study draws upon Lockwood's (2013) model of students' combinatorial thinking (Figure 1), which proposes three basic components of students' counting (*expressions/formulas, counting processes*, and *sets of outcomes*) and elaborates on the relationships between these components. Lockwood (2013) defines the set of outcomes as the "collection of objects being counted – those sets of elements that one can imagine being generated or enumerated by a counting process" (p. 253). In terms of the model, the idea of systematic listing, especially the act of reflecting on how to create an organized list of outcomes, lies in the relationship between counting processes and sets of outcomes.

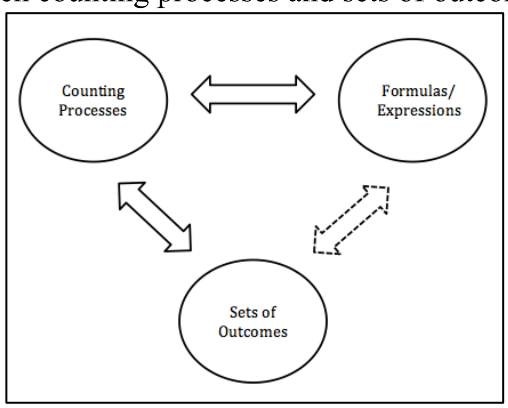


Figure 1 – Lockwood's model of students' combinatorial thinking (p. 253)

A major motivation for the focus on outcomes is that to solve a counting problem correctly, we ultimately must know that we have counted all of the desirable outcomes exactly once. Students, however, can tend to gloss over the outcomes in favor of moving too quickly to formulas or techniques to solve counting problems (e.g. Kavousian, 2006; Lockwood, 2011). Although the focus on sets of outcomes stems primarily from Lockwood's (2011, 2013) previous work, other researchers (English, 1991, 2005; Hadar & Hadass, 1981; Polaki, 2005; Shaughnessy, 1977) have acknowledged that emphasizing the set of outcomes could support counting activity. In this paper, our premise is that students may benefit from an explicit focus on sets of outcomes, and our work is motivated by a broader goal of investigating whether (and if so, how) students may benefit from work with sets of outcomes.

Listing Strategies. A significant challenge with solving counting problems is that it can be difficult to convince oneself that all of the desirable outcomes have been counted exactly once. Constructing a systematic, organized list can allow us to make convincing arguments about why we have counted all of the outcomes. Some researchers (e.g., English, 1991; Halani, 2012) have discussed listing in combinatorial tasks across a number of age levels, and a common and effective listing strategy is an *odometer* strategy. English identified combinatorial strategies in her work with young children (age 4 to 9 years), describing the odometer strategy being sophisticated and defining it as having a consistent and complete cyclical pattern with "a 'constant' or 'pivotal' item…Upon exhaustion (or apparent exhaustion) of the item, a new constant item is chosen and the process repeated" (p. 460). Building on English's work, Halani (2012) identified an *odometer* way of thinking. A major benefit of the odometer strategy is that it convincingly provides a rationale for why no outcome is missed. We hypothesize that systematic listing can give students a mechanism by which to convince themselves that they have all of the outcomes – something that is not trivial.

The section above is meant to highlight that while there is some mention of outcomes in the combinatorics education literature, the treatment of outcomes is largely implicit. That is, researchers have not set out to systematically test the effectiveness of either students' engagement with outcomes when counting or on instructional interventions that foster such engagement. Similarly, while listing has been identified as a common strategy among students, studies have not targeted the effects of listing on combinatorial performance. In this study, we explicitly study outcomes as a factor that might affect students' success in counting. The study looks beyond describing and categorizing listing strategies in two ways – first, by quantitatively reporting the effectiveness of listing on students' performances on counting problems, and second, by detailing the nature of productive versus unproductive lists in undergraduates' listing strategies. In a recent plenary address, Weber (2013) advocated for an increase in quantitative studies to complement qualitative studies in mathematics education. This study is attempt to respond to this call by balancing quantitative and qualitative results.

Methods

Participants and data collection. Forty-two undergraduate students participated in the study. These students were enrolled in an introductory psychology course at a large Midwestern university, and they received extra credit for their participation. Demographic information revealed varying degrees of experience with counting problems and suggested that almost all of the students had seen counting problems before (most typically in high school but not formally in college). While the psychology students were novice counters, counting problems do not require any mathematical prerequisites to solve and do not preclude such students from being able to approach the problems. The students completed a written assessment consisting of counting tasks, which took about 60 minutes to complete. We administered the assessments to students in three iterations on three different days. In a given day, the students took the survey in groups of 1-6 students at a time. The number of participants for the three days was 13, 19, and 10, and minor adjustments (discussed below) were made between each of the days.

Within each of the iterations, the data collection process was the same. The students were randomly assigned to either a *listing* or a *non-listing* condition, and each assessment involved pre-intervention, intervention, and post-intervention tasks. The tasks were the same for each condition; the only difference was in the prompts given to students. For the *non-listing condition*, students were given written instructions that simply asked them to solve the problems and show their work. For the *listing condition*, the students were also prompted with the following written statement prior to the intervention tasks: "In the following 3 problems, please make an attempt to create a list of what you are trying to count" and were given a verbal prompt, "On these problems, please first make an attempt to list out what you are trying to count."

Tasks. The written assessment consisted of 10 counting problems (12 in iteration 1) that would be accessible to novices, involving relatively simple applications of addition and multiplication. We chose tasks with a variety of sizes of sets of outcomes (some which could not easily be listed by hand) in order to allow us to see if even partial listing might help students count successfully. In all of the problems, even ones in which listing all outcomes was not plausible, we hoped that students would be able to write down outcomes and perhaps use that listing to determine a useful pattern or structure. The statement of each task, its pre-, intervention, or post- status, and the cardinality of the answer, are outlined in Tables 1, 2, and 3. We used Livescribe pens to collect the data, which have technology that allows for written responses and audio to be recorded in real time. The written work is then embedded into a pdf file, and one can "play" the pdf to see what was written in real time. We did not analyze the audio-recordings, as the students sat in the room quietly and completed the written assessment.

Table 1 – Pre-Intervention tasks

Pre-Intervention Task		Answer
Test Questions	How many ways are there of answering an 8-question multiple choice test if there are four possible choices for each question?	65,536
Language Books	There are five different Spanish books, six different French books, and eight different Russian books. How many ways are there to pick a pair of books that are not both in the same language?	118
Committee*	Fred, Jack, Penny, Sue, Bill, Kristi, and Martin all volunteered to serve on a class committee. The committee only needs 3 people. How many committees could be formed from the 7 volunteers?	35

Table 2 – Intervention tasks

Intervention Task		Answer
Apples and Oranges	You have 8 identical apples and 8 identical oranges. You need to take some of this fruit to a friend's house, and you don't want to show up empty-handed (you must bring at least 1 piece of fruit). How many possibilities are there for what fruit you could bring?	80
Dominos	A domino is a rectangular tile that has a line dividing one side into two halves. There can be dots on each half, ranging in number from 0 to 6. If you had to make a complete set of dominos, how many dominos would you have to make?	28
Lollipops	You want to give 3 identical lollipops to 6 children. How many ways could the lollipops be distributed if no child can have more than one lollipop?	20

Table 3 – Post-Intervention tasks and justification
(*The Committee and ABCZZZZ tasks were interchanged between Iterations 2 and 3)

Post-Intervention Task		Answer
Cards	In a standard 52-card deck there are 4 suits (Hearts, Diamonds, Spades, Clubs), with 13 cards per suit. There are 3 face cards in each suit (Jack, Queen, King). How many ways are there to pick two different cards from a standard 52-card deck such that the first card is a face card and the second card is a Heart?	153
ABCZZZZ*	How many arrangements are there of the letters A, B, C, Z, Z, Z, Z, where the A, B, and C occur alphabetically (they do not have to appear together as a group)?	35
3-Letter Sequences	You want to make a 3-letter sequence of using the letters a, b, c, d, e, and f. Letters may be repeated and the sequence must contain the letter e. How many such 3-letter sequences are there?	91
CATTLE	How many arrangements of the word CATTLE have the two T's appearing together either at the beginning or the end of the word?	48

Adjustments made between iterations. After Iteration 1, a few minor adjustments were made. In an initial analysis of Iteration 1, we shortened the survey by removing two problems involving non-consecutivity, which seemed to be too challenging. Based on the first author's prior experience with the CATTLE problem, and after having seen students' work in Iteration 1, we felt that a problem involving license plates may not provide insight into listing as much as the CATTLE problem, with which we replaced it. Finally, we made minor adjustments to clarify the wording of the Apples and Oranges and Domino problems. We preserved the same tasks as intervention tasks in all three iterations. Between Iterations 2 and 3, we interchanged the Committee and the ABCZZZZ as pre-intervention and post-intervention tasks in an attempt to mitigate order effects. Ultimately, we do not see these small adjustments as being problematic for our overall findings from the study. As the Results section show, our findings focus on the effects of listing, regardless of condition. While one could argue that making these changes might have helped students perform better (by providing easier tasks, say), we are still able to examine the students' listing behavior and its effect on their solving of counting problems.

Analysis. Initially, the first author coded the responses according to correctness (*correct* or *incorrect*) and we analyzed effects of condition on correctness. As will be described in the Results section, we were led to conduct additional analysis on students' listing behavior. The first author then coded the student responses according to four categories of listing: *no listing*, *articulation*, *partial listing*, and *complete listing*, respectively. A code of *no listing* was given if there was no attempt at any kind of partial or complete list. Typically a student who did not list wrote a numerical value or some kind of formula or expression. A code of *articulation* emerged during analysis, as some responses involved more than only providing a formula or numerical answer, but they were not suggestive of even a partial list. This articulation code was given when a student wrote down at least one outcome but did not actually create any kind of list. A code of *partial listing* was given if the student created a partial list of the outcomes but did not write the entire list correctly or truncated their listing when they identified a pattern. A code of *complete listing* meant a student provided a complete, correct list of the outcomes. All problems were coded one problem at a time to maximize the consistency in coding per problem. These listing codes were coded independently of the listing/no-listing intervention conditions, with the aim of establishing whether the students had listed at all on a problem (regardless of condition).

The quantitative findings suggested that listing could be potentially beneficial for students (discussed in the Results section), and we were thus motivated to look more closely at students' work to learn more about what aspects of listing might be particularly helpful, and why. For the qualitative analysis, then, we reviewed the pdfs of the students' work and watched back through the real-time work, focusing especially on those solutions that had been coded as correct and involving partial listing. We used the constant comparison method (Strauss & Corbin, 1998) to document features of lists that yielded correct versus incorrect results, and for each student's work we recorded phenomena that shed light on the nature of productive listing. The goal was to characterize which aspects of students' listing behavior contributed to successful solutions.

Results

In this section we present both quantitative and qualitative results. Together these contribute to the overall narrative that certain listing behaviors and activities seem to be beneficial for students' counting, and that listing warrants more attention in combinatorics education research.

Quantitative Results

In the following analysis, only problems where the answer was clearly correct or incorrect, and where the listing behavior was clear were used (a total of 352 problems – some problems were excluded because of poor Livescribe pen capture). While we had slightly changed some tasks between iterations, as described above, no significant difference in either number of items answered correctly or the listing behavior was found between iterations. Figures 2 and 3 show the mean and standard error of the number of problems correct and number of problems listed on for each condition in each phase of the experiment.

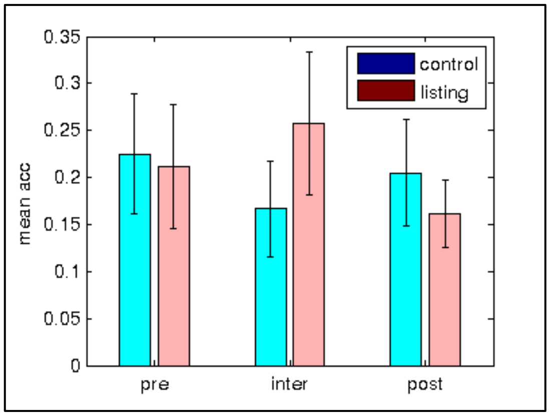


Figure 2 – Mean correct by condition

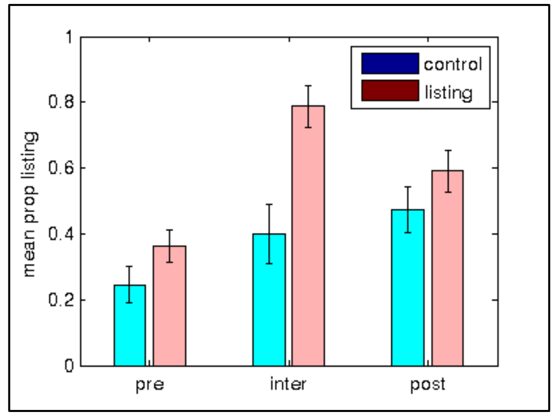


Figure 3 – Mean proportion listing by condition

To address Research Question 1 (*Does engaging in systematic listing have a significant effect on students' solving counting problems correctly?*), we measure student performance by number of questions answered correctly. On the whole, students struggled to solve these problems correctly, with only 24% (84/352) accuracy overall. We also found that student performance on post-test questions did not differ significantly between conditions. That is, students in the *listing condition*, who were instructed to list, did not perform significantly better than students in the *non-listing condition*, who were not explicitly instructed to list. In other words, simply being instructed to list did not have a significant effect on students' performance. Additionally, if we look at the difference between the mean number of questions answered correctly in the pre-intervention tasks versus the post-intervention tasks, we again do not see any significant difference between conditions. In sum, then, the intervention, when measured by condition, did not have an effect on students' performance.

However, during the study we did notice many instances of students in the non-listing condition actually engaging in listing as they solved problems, and vice versa. Therefore, we decided to examine not the effect of a listing/no-listing condition, but rather to explore the effects of students' actual listing behavior. Indeed, if we looked at listing behavior itself, we discovered that listing had an overall positive effect on correctly solving a problem (here, we take *listing* as including a code of either partial or complete listing). We performed two tests to confirm this.

In the first test, we asked *Were students just as likely to get problems correct, regardless of whether they listed or did not list?* To answer this we calculated, for each student, the proportion of problems on which they were *correct and listed* out of those problems on which they *listed*, as well as the proportion of problems on which they were *correct and did not list* out of the all problems on which they *did not list*. This of course required looking only at those students who both listed and did not list (n=39). Performing a paired t-test we found a significant difference between these two proportions ($t(37) = 3.92$, $p < .00038$). This is evidence for saying that there is a correlation between accuracy and listing behavior. Figure 4 shows a summary of these proportions, averaged across students. The first column represents average number of problems on which students listed, with the top of the column being the mean proportion of problems where students listed and got the problem correct (mean = 0.29, SD = 0.28). The second column represents the same mean values but for problems where students did not list (mean = 0.11, SD = 0.21). The significance we show here is the difference between the relative size of the top portions of these columns to the entire column.

In the second test, we asked a similar question: *Were students just as likely to have listed on problems they got correct as those they got incorrect?* For each student (who had both correct and incorrect answers *n*=29) we calculated a proportion of number of problems *with listing and correct* out of *number correct*, and number of problems *with listing and incorrect* out of *number incorrect*. Applying a paired *t*-test we again find a significant difference ($t(29) = 5.32$, $p < .000011$). Summarizing across students we find an average proportion of listing and correct to correct (mean = 0.69, SD = 0.38), and an average proportion of listing and incorrect to incorrect (mean = 0.40, SD = 0.26). Figure 5 summarizes these results – again, the important feature is the large difference in the relative size of the top portion of the column to the entire column.

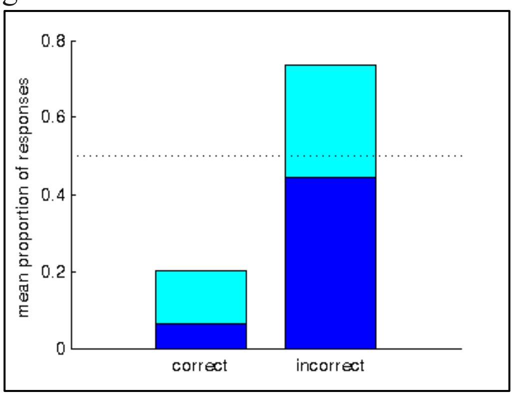


Figure 4 – First paired *t*-test

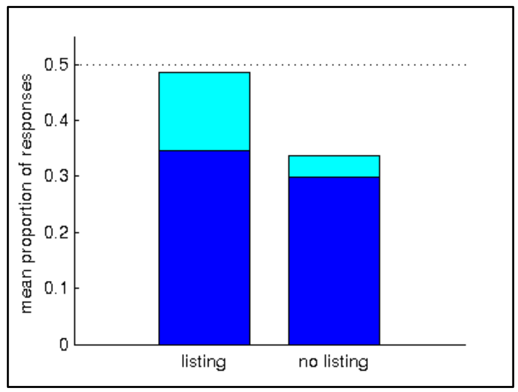


Figure 5 – Second paired *t*-test

In sum, the quantitative results show that while no difference was found among condition, and while students' actual listing behavior was not necessarily influenced by the instructional intervention in the listing condition, students' actual listing behavior was positively correlated with correctly answering counting problems. We note that while there is a correlation between listing and correctly answering a problem, we do not claim causation. We acknowledge that it may be the case that stronger students may naturally list, and that is why we see the positive correlation. However, regardless of whether success leads to listing or vice versa, the correlation is promising – if the more successful counters are listing, perhaps listing deserves more attention as a pedagogical focus. Given students' clearly documented and sustained struggles with counting problems, these initial quantitative findings suggest that listing may be a valuable way to help students count more successfully. We feel that the findings at least warrant more attention, particularly because our results also suggest that instructional intervention was not consistently effective in getting students to list. We were therefore motivated to study the listing that students did in more detail, which we did by examining the students' written work.

Qualitative results

These qualitative results stemmed from the basic quantitative findings, as discussed above, with the aim of answering Research Question 2 (*What are features of productive or unproductive* *lists that students generate, and how does listing activity differ in the generation of productive versus unproductive lists?*). In our analysis, we distinguish between *productive* lists and *unproductive* lists. We take productive lists to mean any lists, partial or complete, which were generated on a problem that the student solved correctly. Unproductive lists are lists that were generated on a problem that was incorrect. Because of the nature of our data, we cannot make conclusive statements about whether or not a particular list actually caused a student to answer a problem correctly. However, for analytic purposes we found the productive versus unproductive distinction to be helpful as we tried to determine potential aspects of listing that seemed particularly beneficial for students' counting. Below, we first discuss features of productive lists (providing contrasting examples of unproductive lists), and then we present additional noteworthy aspects of listing that arose among multiple students. These qualitative results complement the quantitative results presented previously, helping to paint a clearer picture of precise ways in which listing seemed to be effective for students in some situations.

Features of productive lists

In this section we discuss three key features of productive lists, which, while not necessarily present in every productive list, are representative of overall characteristics of productive lists.

Useful notation and appropriate modeling of outcomes. Students who wrote productive lists typically found a suitable notation that appropriately modeled an outcome. In some problems, when an outcome is fairly self-evident and can easily be written on the page literally (such as letter or number sequences), figuring out a meaningful notation may be trivial. Other problems may require extra work to translate the outcome into something that can be written down and listed, and many problems require some translation of the problem into a useful notation. The Lollipop problem highlights the value of a usable notation and shows how students displayed a variety of notations even on the same problem. For example, Student 331[1] let the letters A, B, C, D, E, and F represent the six students. Her outcomes were sets of 3 letters (representing 3 students), which she listed lexicographically (Figure 6).

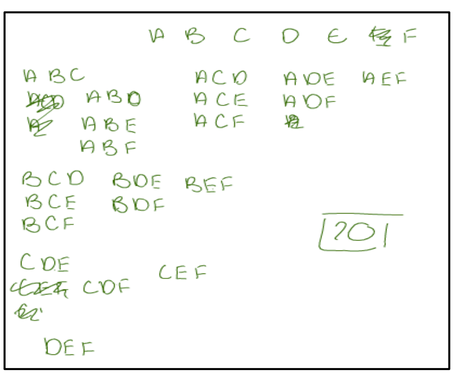


Figure 6 – Student 331's Lollipop problem

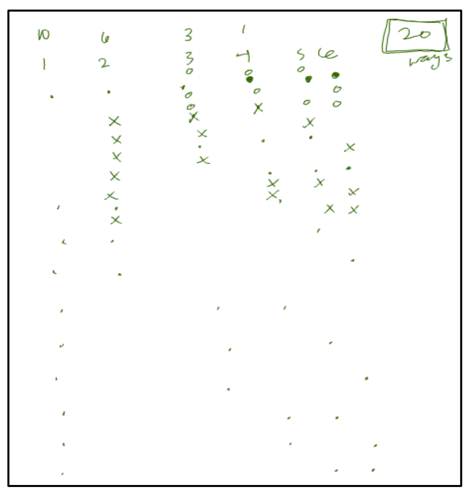


Figure 7 – Student 342's Lollipop problem

Student 342 encoded the outcomes by creating a table that labeled columns 1-6 for the students (Figure 5). The student represented an outcome as a row of three marks, with one mark

---

[1] Because gender information was not collected for each participant, we will refer to students whose number sums to an even number (such as 121) as male and an odd number as female. As a matter of interest, students with numbers ending in 1 were in the listing condition, and students with numbers ending in 2 were in the non-listing condition.

in exactly one column, and the total number of rows gave the final answer. This table is particularly effective in its use of different marks for rows that have marks in different first columns. This notation allowed for her to count up the rows, but it provides more information than only the final numerical answer.

In contrast to these lists that show useful notations that facilitated a way to model outcomes, some students' notations seemed problematic. For example, in Student's 241 work on the Lollipops problem (Figure 8), her labeling of Lollipops 1, 2, and 3 suggests that she was thinking of the lollipops as distinct. When she writes permutations of the numbers 1, 2, and 3 beneath six children, this suggests that she has not clearly articulated what constitutes an outcome. Her issue may not be merely one of notation – she may have some incorrect notion of what the problem is asking, etc. – but her lack of a clear notation certainly does not help her on this problem. The notation did not facilitate a correct articulation of what constituted a desirable outcome. In sum, a key aspect of being able to create a productive list is to correctly model the outcomes, which often involves developing an efficient notation.

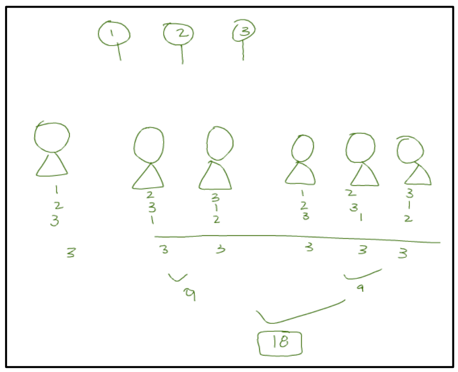


Figure 8 – Student 241's Lollipop problem

Organized strategy. Another feature of productive lists was that they often seemed to be developed with an intentional organizational strategy. Student 331's work on the Lollipop problem (Figure 6 above) exemplifies the odometer strategy. In order to list the set of 3 letters, she began by holding the first and seconds element constant, and then cycling through the last elements in the order of how she initially wrote the six letters that represented the students. Once she had similarly cycled through each possibility for the second letter, she could move to the next choice for the first option and repeat the process. Because we could watch her make the list in real time, we know that she did in fact implement this process as she listed outcomes.

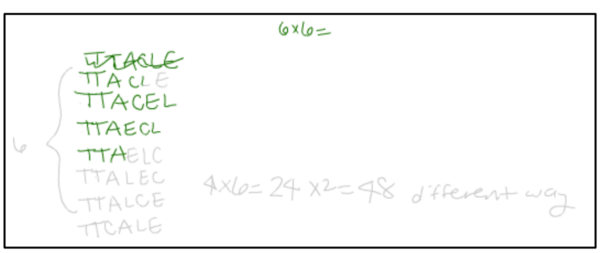


Figure 9a – Student 131's partial list of the CATTLE problem

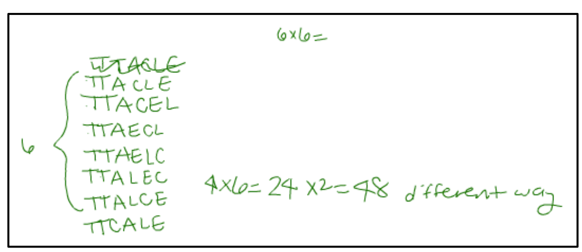


Figure 9b – Student 131's complete work on the CATTLE problem

A similar organizational strategy was exemplified by a number of students in the CATTLE problem. These students did not create an entire list for this problem; instead, they wrote out some of the arrangements of the letters C, A, L, and E, identified a pattern, and used multiplication to calculate the total. For example, Student 131 wrote out the two Ts and arrangements of the letters A, C, L, and E. The student's real-time listing shows an attempt on his part to remain organized and systematic. He first wrote TTACLE but then crossed it out, and we infer from the rest of his work that he sought to list alphabetically. Figure 9a shows he then proceeded to write the first alphabetical outcome, TTACEL, followed by TTAECL. As he was writing TTAELC, he seemed to realize that he had missed another outcome starting with "AC," and so he went back and added TTACLE to the top of his list, pairing it with TTACEL. Figure 6a shows him in the process of going back and adding TTACLE to the top of the list. He then proceeded to complete an alphabetical list of arrangements starting with A. He wrote one of the arrangements starting with C (TTCALE) but then seemed to notice a pattern. Figure 9b shows his final list, in which he noted there were 6 options for the starting letter, and he multiplied this 6 by 4 and then by 2 to yield the correct answer. This example shows a student using an organized strategy on a partial (as opposed to a complete) list that ends up being productive. The student was intentionally organized in her listing, to the point of going back and adding an outcome where it best belonged within his scheme. This organized, near alphabetical pairing of certain outcomes helped ensure that he had all of the outcomes.

In contrast to the organized lists above, some unproductive lists lacked the kind of organizational strategy that could easily account for all of the outcomes. For example, on the ABCZZZZ problem (Figure 10), Student 661 correctly wrote out a number of outcomes. There is some initial organization, as she cycles the ABC through the Zs. However, beyond that the student is not systematic in her listing, and as a result many of the outcomes are missed.

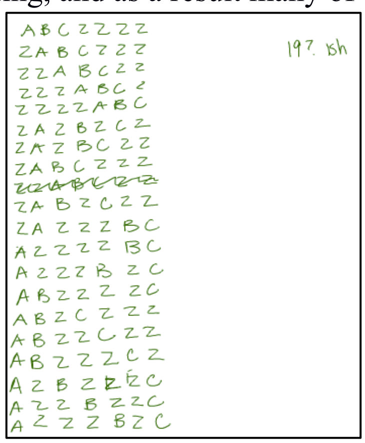


Figure 10 – Student 661's ABCZZZ problem

Evident Structure. Some productive lists had an obvious structure that elicited a certain way of organizing the outcomes. As noted in the literature review, a potential benefit of listing is that, if done carefully and systematically, listing can provide concrete evidence that all of the outcomes have been accounted for. The structure evident in some students' lists contributed to a convincing argument that all of the outcomes were being counted exactly once.

Student 252's work on the Lollipop problem (Figure 11) clearly yielded a list that suggests a particular sum: (5+4+3+2+1)+(4+3+2+1)+(3+2+1)+(2+1)+1. The student encoded triangles as the students (distinct because they are in a line), with rows of circles under three of the triangles representing an outcome. She wrote down the outcomes systematically by first holding constant the circles in the first two columns, and then cycling the third through the remaining columns. Then, while still keeping the first entry static, she moved the second circle to the second triangle and cycled through the all of the possibilities for the first triangle, continuing in this way while keeping the first item constant (apparently making use of the odometer strategy). The first sum of 4+3+2+1, then, includes all the outcomes with a circle under the first triangle (or, with the first child receiving a lollipop). She then proceeded to move the first circle to the second triangle and repeated the process, yielding 3+2+1. She continued to repeat the process to produce the remaining sums. The structure of the sum visually pops out of the list, making apparent how the student meaningfully organized the outcomes to count them effectively.

In contrast to Student 252's work, on the same Lollipop problem Student 442 wrote the 20 outcomes in a 5*4 array (Figure 12). Given the student's subsequent writing of the binomial coefficient for "6 choose 3," it may be the case that he had already guessed or arrived at the answer and wanted the array to reflect the answer of 20. However, the point is that the array, while correctly representing an answer of 20, does not offer much insight into the structure of the list or why the student may be convinced that all of the outcomes are counted. Unlike the work in Figure 11, there is no further insight gained by the structure of how the list is written. In fact, while on some problems an array might provide some insight, here the particular arrangement of the two outcomes in the array hides any relevant structure of the set of outcomes.

Creating a list that highlights a particular structure is an effective way of connecting a counting process to a set of outcomes, and examining the set of outcomes can be an important means by which to be sure a student is counting correctly. A list with a transparent structure may provide concrete evidence that the list may be correct. Making such a list may be a productive verification strategy that should be investigated more thoroughly in subsequent studies.

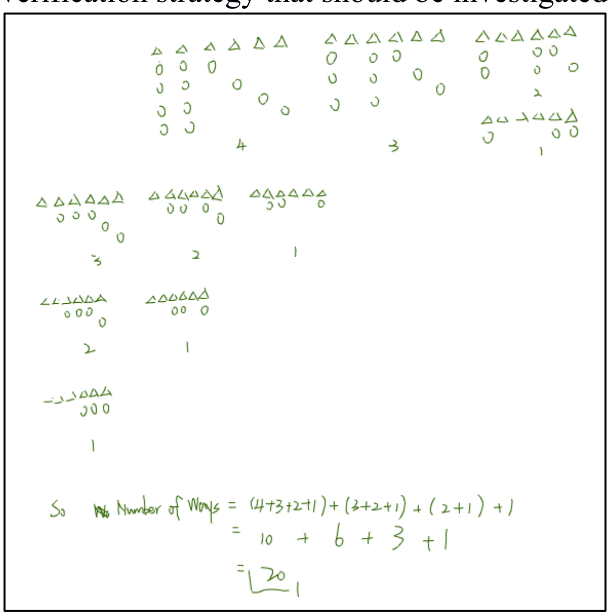


Figure 11 – Student 252's Lollipop

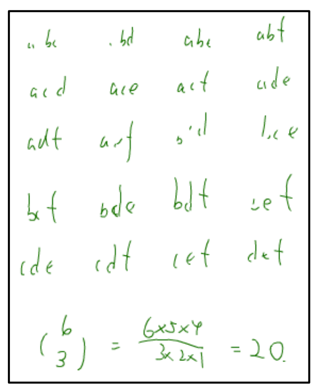


Figure 12 – Student 442's Lollipop problem

Other insights into productive listing

In addition to identifying features of productive lists, we also identified two themes across students' work that shed light on productive listing: creating productive lists seemed to affect students' work on other problems, and even partial listing proved to be beneficial.

Productive listing experience seemed to affect students' work on other problems. Perhaps one of the more surprising results that came out of the qualitative analysis was to see the dynamic way in which some students' work unfolded across problems. The real-time pdfs showed that students were much more creative and dynamic in their listing than can be seen simply from the written work on the page. This came out most pointedly as students moved back and forth between problems. On several occasions it seemed that successful listing on one problem led students to go back and revisit previous problems, incorporating listing strategies to arrive at correct answers. As an example of this, Student 121 worked on the Domino problem, initially drawing out dominos and listing 0-0 through 0-6 and writing *6=36 (see the dark green but not grayed out writing in Figure 13a). The grayed out work would suggest that he then proceeded to list the rest of the dominos, cross out the duplicates, and arrive at the correct answer of 28. However, immediately following writing down 36, he moved on to the next problem. The dynamic recording reveals what was, to us, a surprising phenomenon.

Student 121 then solved the Lollipop problem, engaging in very careful and systematic listing, so that she ultimately arrived at the correct answer of 20 using a very well organized list. Upon completing the Lollipop problem, he immediately returned to the Domino problem, subsequently systematically listing all of the dominos, crossing out duplicates and arriving at the correct answer of 28 instead of 36 (Figure 13b). While we cannot know for sure his progression of strategies, because we could not ask follow up questions, it is interesting that directly following successful, systematic listing on the Lollipop problem, he used listing to fix an initially incorrect answer. In this case, we conjecture that the student gained an important insight as he listed in the Lollipop problem – namely, the Lollipop problem does not count rearrangements of the same three students as distinct outcomes. By listing out the Lollipop possibilities, he knew that he did not want to count sets {1, 2, 3} and {2, 1, 3} distinctly toward the total. The same is true of dominos – domino 1,2 is the same as domino 2,1 – and his subsequent behavior on the Domino problem suggests that the work on the Lollipop problem led him to make a change in her strategy on the Domino problem. One could argue that it was not the act of listing itself that caused his realization, but rather that it might be caused by exposure to another problem with outcomes of a similar kind. However, we contend that the listing on the Lollipop problem drew attention to the nature of outcomes in ways that simply trying to solve the problem without listing might not have done. Additionally, the student's subsequent systematic and detailed list on the Domino problem suggests that the act of listing itself was something he carried over from the Lollipop problem and chose to utilize in his subsequent solution of the Domino problem.

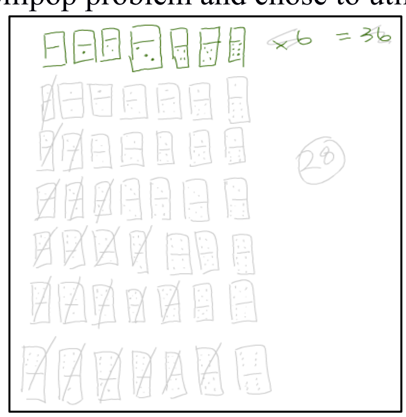


Figure 13a – Student 121's initial Domino

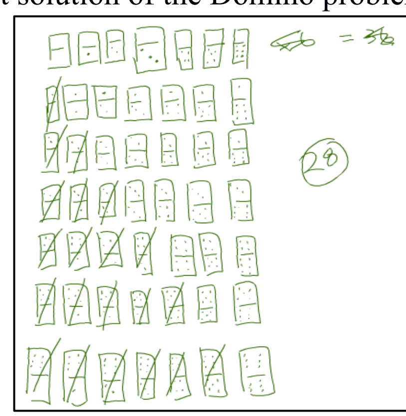


Figure 13b – Student 121's correct Domino

Other students had similar trajectories in which they revisited a problem after what they perceived as successful listing on another problem, suggesting that Student 121's case was not a one-time phenomenon. We feel that the effect of successful listing on other problems is something that could be investigated more explicitly in further studies.

Even partial listing can be productive. The second theme we identified was that students could, at times, productively arrive at the correct answer by creating only a partial list without having to list all of the outcomes completely. Often, some type of encapsulation process was observed in the list, and while students at times overgeneralized, many students were able to identify and correctly use a pattern they saw in a partial list. Student 131's work on the CATTLE problem above is one such example of a productive partial list.

As another example, Student 431's work on the Apples and Oranges problem demonstrates a progressive streamlining process that emerged during the students' attempts to solve. Figure 14 shows the student began a fairly detailed and complete list. Then, we see an increasingly streamlined listing process, and by the end the process of listing each outcome is encapsulated and truncated. The student is organized, displays a structure in the list, and he arrives at the correct solution without having actually listed all 80 outcomes.

If listing were only beneficial if the problem's solution can easily be listed completely, the value of listing would have serious limitations, as most counting problems have solutions that cannot easily be physically listed by hand. Counting problems regularly have very large sets of outcomes, and it is unrealistic to claim that students might be able to list all desirable outcomes as they solve counting problems. The fact, then, that we have evidence that listing appears to be a useful strategy even on problems which students may not choose to (or be able to) create complete lists is promising for helping students be successful in a variety of counting situations.

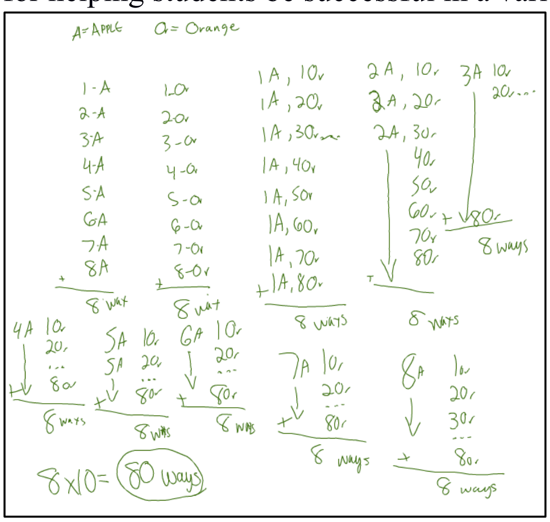


Figure 14 – Student 431's Apples and Oranges problem

Discussion and Implications

We briefly mention several points of discussion. First, the number of instances in which students attempted to list, as well as whether such listing was productive, varied from problem to problem. Certain problems did seem to elicit listing (partial or complete) more than others, such as when comparing the CATTLE (more productive listing) and ABCZZZZ (less productive listing) problems. While the role of a specific problem in students' listing warrants further investigation, there may already be some implications for instruction. If we indeed accept the premise that it might be worthwhile for students to do at least some listing as they learn counting initially, then it is noteworthy that teachers should take into account what problems might best facilitate listing for students, and at least be aware that not all problems can are equally effective in eliciting listing. From a research perspective, the effects of certain problem features and problem types on students' listing behavior warrant further study.

Additionally, a noteworthy factor in helping students to list productively seems related to having a careful disposition toward mathematical work (in this case, articulating outcomes and listing), which is in line with findings other researchers have shared (e.g., Hadar & Hadass, 1981; Lockwood, in press a). We suspect that students who failed to implement organizational strategies were not always making a mathematical error, but rather that they were, at times, not being careful and deliberate in their work. This is an aspect of listing, and counting more generally, that needs to be investigated further, perhaps by explicitly examining how metacognitive aspects of problem solving affect students' counting.

We also observe that the results support the notion that a focus on sets of outcomes is important, as Lockwood (2013) has proposed. Systematic listing orients students with what constitutes a desirable outcome, ensuring that they understand what they are trying to count. Identifying structure and organizational techniques in a list reinforces productive counting processes that can help student generate patterns and avoid overcounting. The value of listing evidenced in this study validates further work on sets of outcomes so that we might better understand ways in which outcomes might productively be used to help students count.

Finally, our minimal intervention of prompting students to list via a simple written and verbal prompt was not enough to cause them to list consistently. More work needs to be done to investigate alternative methods of instruction to help students gain experience with and to appreciate the benefits of systematic listing. Especially given the fact that in some cases listing on one problem affected students' work on other problems, it seems promising to target whether or not (and if so, how) instructional interventions can be designed to help students develop listing behavior. Specifically, we can investigate how to help students generate and identify meaningful patterns in lists. Additionally, it may be instructive for students to see productive versus unproductive lists in others' work, in order to reflect on how helpful lists might be generated.

In conclusion, our aim in this study was to examine whether or not having students systematically list might be a potentially helpful intervention in their solving of counting problems. Our findings suggest that, while our particular intervention was not entirely effective as intended, the listing (and even partial listing) of outcomes was positively correlated with students' correct solving of counting problems. The qualitative analysis gave further insight into features of productive lists that suggest aspects of listing that could be incorporated into instructional interventions. In light of clear evidence that students at all levels struggle with correctly solving counting problems, we have uncovered one factor that is a significant factor in successful counting – having undergraduate students engage in systematic listing of what they were trying to count. These results also support prior work by Lockwood (2011, 2013) that points to the importance of focusing on sets of outcomes. Our findings indicate that more needs to be done in order to develop instructional interventions that will facilitate listing, but they also suggest that such an endeavor holds much promise for improving students' counting.

References

Annin, S. A., & Lai, K. S. (2010). Common errors in counting problems. *Mathematics Teacher, 103*(6), 402-409.

Batanero, C., Navarro-Pelayo, V., & Godino, J. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*, 181-199.

Eizenberg, M. M., & Zaslavsky, O. (2004). Students’ verification strategies for combinatorial problems. *Mathematical Thinking and Learning, 6*(1), 15-36.

English, L. D. (1991). Young children's combinatorics strategies. *Educational Studies in Mathematics, 22*, 451-47

English, L. D. (2005). Combinatorics and the development of children’s combinatorial reasoning. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (Vol. 40, pp. 121-141): Kluwer Academic Publishers.

Godino, J., Batanero, C., & Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60*, 3-36.

Hadar, N., & Hadass, R. (1981). The road to solve combinatorial problems is strewn with pitfalls. *Educational Studies in Mathematics, 12,* 435-443.

Halani, A. (2012). Students’ ways of thinking about enumerative combinatorics solution sets: The odometer category. In the *Electronic Proceedings for the Fifteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education.* Portland, OR: Portland State University.

Kapur, J. N. (1970). Combinatorial analysis and school mathematics. *Educational Studies in Mathematics, 3*(1), 111-127.

Kavousian, S. (2008) *Enquiries into undergraduate students’ understanding of combinatorial structures.* Unpublished doctoral dissertation, Simon Fraser University – Vancouver, BC.

Lockwood, E. (2011). *Student Approaches to Combinatorial Enumeration: The Role of Set-Oriented Thinking.* Unpublished doctoral dissertation, Portland State University.

Lockwood, E. (2013). A model of students’ combinatorial thinking. *Journal of Mathematical Behavior, 32,* 251-265. Doi: 10.1016/j.jmathb.2013.02.008.

Lockwood, E. (In press a). The strategy of solving smaller, simpler problems in the context of combinatorial enumeration. To appear in *Research in Collegiate Mathematics Education*.

Maher, C. A., Powell, A. B., & Uptegrove, E. B. (Eds.). (2011). *Combinatorics and Reasoning: Representing, Justifying, and Building Isomorphisms.* New York: Springer.

Martin, G. E. (2001). *The Art of Enumerative Combinatorics*. New York: Springer.

Mellinger, K. E. (2004). Ordering elements and subsets: Examples for student understanding. *Mathematics and Computer Education, 38*(3), 333-337.

Polaki, M. V. (2005). Dealing with compound events. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 191-214). New York: Springer.

Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small-group, activity-based, model building approach to introductory probability at the college level. *Educational Studies in Mathematics, 8,* 295-316.

Strauss, A. & Corbin, J. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.

Weber, K. (2013). Benefits and limitations of complementing qualitative research with quantitative studies: Lessons learned from four studies on proof reading. Presented at a plenary session of the *16th Conference on Research in Undergraduate Mathematics Education.* Denver, Colorado.

EXAMINING STUDENTS' COMBINATORIAL THINKING THROUGH REINVENTION OF BASIC COUNTING FORMULAS

Elise Lockwood — Oregon State University
Craig Swinyard — University of Portland
John S. Caughman, IV — Portland State University

Counting problems provide an accessible context for rich mathematical thinking, yet they can be surprisingly difficult for students. To foster conceptual understanding that is grounded in students' thinking, we engaged a pair of undergraduate students in a ten-session teaching experiment. The students successfully reinvented four basic counting formulas, but their work revealed a number of unexpected issues concerning justification in counting. In this paper, we describe the students' successful reinvention of the four counting formulas, we critically examine their combinatorial reasoning in terms of Lockwood's (2013) initial model of students' combinatorial thinking, and we offer several directions for further research.

Key Words: Combinatorics, Reinvention, Counting Problems, Teaching Experiment

Introduction and Motivation

Enumerative combinatorics (the solving of counting problems) has applications in probability and computer science, and its accessible yet challenging problems provide a rich context for developing students' mathematical justification and problem solving skills. As a result, counting problems have gained traction in K-12 and undergraduate curricula in recent years, particularly in probability units and in undergraduate discrete mathematics courses. Both researchers (e.g., Hadar & Hadass, 1981; Maher, Powell, & Uptegrove, 2011) and textbook authors (e.g., Martin, 2001; Tucker, 2002) have noted that counting problems facilitate deep and critical mathematical reasoning. For example, Tucker (2002) emphasizes that counting requires "logical reasoning, clever insights, and mathematical modeling" (p. 169). Martin (2001) similarly points out, "One of the things that make elementary counting difficult is that we will encounter very few algorithms. You will have to think." (p. 1). Often one of the first topics a student encounters in an undergraduate discrete mathematics course, counting offers a rich playground for developing the type of mathematical justification skills necessary for advanced courses. However, in spite of the practical applications of counting problems and their potential to foster rich mathematical thinking, student difficulties with counting persist (e.g., Batanero, Godino, & Navarro-Pelayo, 1997; Eizenberg & Zaslavsky, 2004). Research is needed that explicates how students can effectively comprehend basic counting principles.

The aim of our study was to gain insight into how students might come to reason coherently about four basic counting formulas: $n!$, n^r, $\frac{n!}{(n-r)!}$, and $\frac{n!}{(n-r)!\,r!}$. Textbooks typically present these formulas early on, following each with numerous examples, and students are generally expected to apply the formulas in various contexts throughout the remainder of the course. Research (e.g., Batanero, et al., 1997; Lockwood, 2013) indicates, however, that students frequently misapply these formulas, which suggests they may not understand when and why these expressions are to be utilized. Recent studies (Swinyard, 2011; Oehrtman, Swinyard, & Martin, 2014) suggest that students can develop coherent reasoning about mathematical concepts via tasks designed to foster their reinvention of precise *concept definitions* (Tall & Vinner, 1981). These studies have also served as evidence that reinvention can provide researchers a lens through which to gain insight into how students come to understand particular mathematical concepts. With this in mind, we engaged a pair of undergraduates in a ten-session teaching experiment, during which they solved basic counting problems and then subsequently generalized their mathematical activity by reinventing the four basic counting formulas. In this paper, we report on the students' reinvention of the four formulas, addressing the following questions:

1) *How might students reinvent these four basic counting formulas?*
2) *What cognitive issues might arise for students as they reinvent and use these formulas?*

Literature Review

Student difficulties with counting. Researchers have clearly established that students have difficulty with even basic counting tasks. These struggles can be seen by overall low success rates on counting problems. For example, in a study conducted by Eizenberg and Zaslavsky, only 40% (43 of 108) of initial solutions provided by undergraduates were correct (2004, p. 31), and Lockwood similarly found that only 42 out of 103 (41%) problems given to 22 postsecondary mathematics students were answered correctly (2011). Additionally, some researchers have identified factors that might lead to such difficulties, including over-counting and confusion about when order matters (Annin & Lai, 2010; Batanero, et al., 1997; Hadar & Hadass, 1981). Eizenberg & Zaslavsky (2004) also point out the fact that counting problems, with their large numerical answers, can be difficult to verify.

Research on combinatorial concepts. In this study, we had students reinvent formulas for basic combinatorial notions, including permutations and combinations. Our work builds on prior research, including that of Piaget and Inhelder (1975), who posited a fundamental difference between the mental processes that combinations and permutations each respectively require, conjecturing that permutations occur at a more formal thought level than combinations. Fischbein and Gazit (1988), and later Dubois (1984) and Batanero, et al. (1997) also investigated the effects of both implicit combinatorial models and particular combinatorial operations on students' counting. Our study contributes to this prior work by targeting students' conceptualizations of permutations and combinations through reinvention of their formulas. The work is framed within Lockwood's (2013) model of combinatorial thinking, in which she describes and relates three components of students' counting (*formulas/expressions, counting processes, and sets of outcomes*) and argues that students should focus on sets of outcomes as they count. The model is elaborated in the theoretical perspective section, which we discuss after a brief mathematical discussion.

Mathematical Discussion – The Multiplication Principle

The *multiplication principle* (MP) is considered by many (e.g., Martin, 2001; Tucker, 2002) to be a foundational aspect of understanding and justifying counting formulas. Among textbooks there does not seem to be much consensus for a precise statement of the MP. Two statements of the MP are given below, the second of which is a generalization of the first.

- "The Fundamental Principle of Counting: If one task can be completed in m ways and another task can be completed in n ways, then the sequence of the two tasks can be completed in $m \times n$ ways." (Richmond & Richmond, 2009, p. 132)
- "The Multiplication Principle: Suppose a procedure can be broken into m successive (ordered) stages, with r_1 different outcomes in the first stage, r_2 different outcomes in the second stage,…, and r_m different outcomes in the m^{th} stage. If the number of outcomes at

each stage is independent of the choices in previous stages, and *if the composite outcomes are all distinct*, then the total procedure has $r_1 \times r_2 \times \ldots \times r_m$ different composite outcomes" (Tucker, 2002, p. 170, emphasis in original).

We discuss the MP here because it is foundational to basic counting in two key ways. First, it provides some justification for solutions to counting problems. To illustrate this, we consider the following problem:

The Quiz Questions Problem – *On a quiz, there are eight multiple-choice questions, each of which has four possible answers (A, B, C, and D). In how many ways could a student complete this quiz?*

To answer this question, we could make the following argument: "We have four options for how to answer the first question – A, B, C, or D. Then, for any of those possibilities, we again have four options for how to answer the second question, yielding 16 possibilities for answering two questions. By the same argument, we'll have four choices for the third through eighth questions, and we continue to multiply by four for each question." This is illustrated by a partial tree diagram in Figure 1. The MP allows us to explain the general process that justifies why repeated multiplication by four is reasonable without having to list all 4^8 (or 65,536) outcomes.

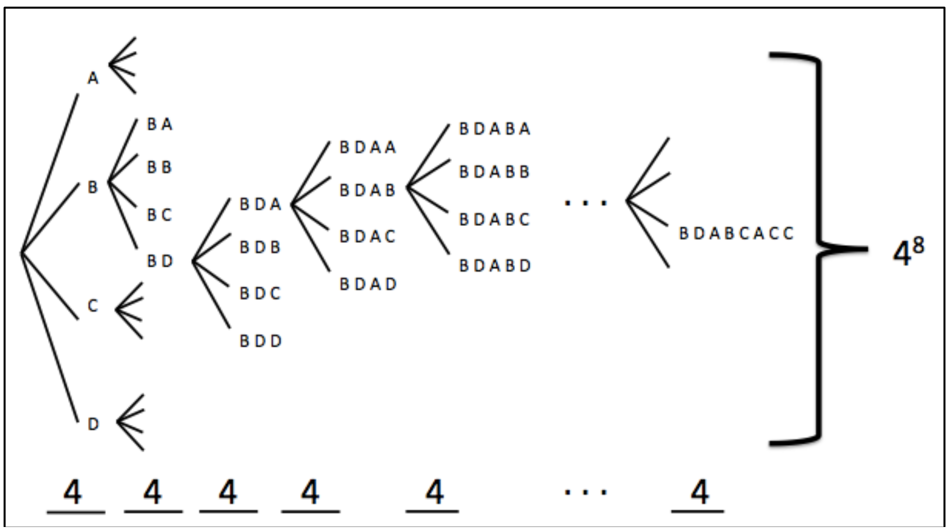


Figure 1 – A tree diagram representing the multiplication principle

Second, the MP is foundational to counting because it underpins the basic counting formulas students encounter. Formulas like $n!$, n^r, $\frac{n!}{(n-r)!}$, and $\frac{n!}{(n-r)!\,r!}$ are each general ways to express particular products that are based on the MP. For example, $n!$ is the number of ways to arrange n objects when objects cannot be repeated. Indeed, if we picture placing objects in distinct, ordered positions, there are n choices for which object goes in the first position, n-1 choices for the second position, and so on. Through the lens of the MP, this can be thought of as distinct, ordered stages, and thus it makes sense that the resulting product of $n!$ gives all of the possible arrangements of n objects. The formula $\frac{n!}{(n-r)!}$ is a variation of the $n!$ formula – instead of arranging all n objects, we may only want to arrange r of n objects. There is again simply a product (based on the MP) that we wish to write, and the formula provides an efficient way to express that product. The n^r and $\dfrac{n!}{(n-r)!\,r!}$ formulas are similarly supported by the MP. Thus, we see that the MP both provides justification for our counting, and it conceptually supports much of the initial counting that students encounter.

Theoretical Perspective

A model of students' combinatorial thinking. Our work is situated within the context of Lockwood's (2013) model of students' combinatorial thinking (Figure 2), which we elaborate here, using the Quiz Questions problem as an example. The model consists of three components – *formulas/expressions*, *counting processes*, and *sets of outcomes*. The *formulas/expressions* are mathematical expressions that yield some numerical value, often considered "the answer" to the counting problem. In the Quiz Questions problem, the expression that gives the final answer is 4^8. The c*ounting processes* are the enumeration processes (or sequence of processes) in which a counter engages as they solve a counting problem. In the Quiz Questions problem, the counting process is the iterative use of the MP to track the options at each stage. The *set of outcomes* refers to the set of elements being generated or enumerated by a counting process. In our example, the set of outcomes consists of the collection of all the ways in which a quiz might be answered, perhaps encoded as sequences of A, B, C, D of length 8. The cardinality of the set of outcomes (65,536 in this case) is equivalent to the problem's solution.

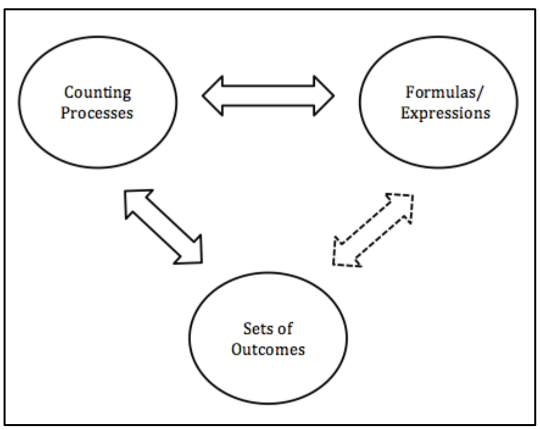


Figure 2 – Lockwood's (2013) model of students' combinatorial thinking

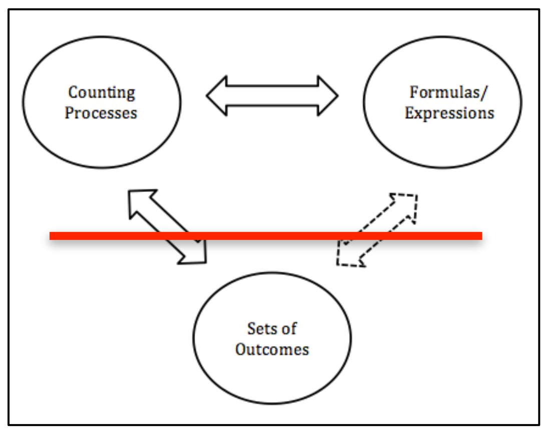


Figure 3 – A depiction of how students typically try to solve counting problems

Lockwood's model also elaborates relationships between the three components and emphasizes the importance of fostering the relationship between counting processes and sets of outcomes. In the initial conceptualization of the model, it was not fully explored what the direct relationship between *formulas/expressions* and *sets of outcomes* might entail, as evidenced by the dotted arrow. Lockwood offered a few conjectural comments, but did not have empirical data to substantiate or articulate the relationship. As such, the model initially served primarily to highlight sets of outcomes as a central construct underlying the counting processes and formulas with which students interact. Additionally, the picture in Figure 3 represents how students typically tend to count – they tend to live and die above the red line. That is, students are inclined not to think about sets of outcomes as they count, but rather they often rely on memorized key words or situations to determine whether order matters or to decide which formula to apply (e.g., Annin & Lai, 2010; Lockwood, 2011, 2013). While some simple problems can be solved without considering outcomes, many commonly troublesome aspects of counting (such as issues of order or reconciling an overcount) can be resolved by students considering outcomes. Despite this, students do not utilize outcomes as often as they could (Lockwood 2011, 2013). The model thus provides language through which to articulate several salient aspects of students' counting activity.

Realistic Mathematics Education. In designing the present study, we conjectured, based on prior research (Lockwood, 2013), that students would have a better chance of successfully using counting formulas (and solving counting problems successfully) if the counting formulas contain meaning for them. Given that our central research goal was to gain insight into how students might reason meaningfully about the four counting formulas previously discussed, we chose not to supply them with these formulas, but rather to have them *reinvent* the formulas (i.e., construct the formulas themselves) by generalizing their work on an initial set of counting problems. We patterned our work after previous studies (Swinyard, 2011; Oehrtman, Swinyard, & Martin, 2014) in which students were able to develop coherent reasoning about mathematical concepts via tasks designed to foster their reinvention of precise mathematical definitions. We thus drew inspiration from the perspective of *developmental research* (Gravemeijer, 1998), which leverages students' informal knowledge and supports them in developing sophisticated, abstract knowledge while maintaining intellectual autonomy (p. 279). In line with Freudenthal's recommendation (1973) to avoid an *antididactic inversion* (where symbolic formalism precedes reasoning), we aimed to create an environment that fosters initial exploration of counting problems that emphasizes sense-making over conventional symbolization.

Methods

The aim of this paired teaching experiment (Steffe & Thompson, 2000) was for two students to reinvent four basic counting formulas through engaging with a variety of counting problems.

Participants. The participants were two above-average students (Thomas and Robin, pseudonyms) who had recently completed an integral calculus course. They were chosen based on the following criteria: 1) they had no formal college-level experience with counting; 2) they demonstrated strong mathematical background and ability; and, 3) they displayed a propensity to engage actively with mathematics and articulate their reasoning. The teaching experiment occurred following Thomas and Robin's freshman and sophomore years, respectively.

Data Collection and Tasks. The experiment consisted of ten 90-minute sessions, which occurred three times a week for about four weeks. The interviews proceeded in three phases, which we describe in detail.

Phase 1 (Sessions 1-4): Solving Initial Counting Problems. Since neither student had prior counting experience, we first engaged them with ten counting problems, with the goal of providing them a common experience from which to generalize their mathematical activity. Three factors played into our task selection. First, we wanted Thomas and Robin primarily to engage in problem solving. Although these students had no formal collegiate experience with counting, we recognized that they may have seen permutation or combinations formulas in high school. To encourage problem solving rather than merely trigger attempts to recall expressions, we wanted to start with problems that do not immediately suggest a direct application of the most basic formulas. Second, we wanted students to use and reason about sets of outcomes, and thus most of the problems involved numbers small enough to facilitate explicit listing. Finally, we aimed for a variety of problems to require a range of approaches. This included problems that might naturally be broken into cases, problems for which the multiplication principle might naturally be applied, problems that might be susceptible to overcounting, and at least a few problems with larger sets of outcomes. We provide a sample of problems from Phase 1 in Table 1, focusing on those that we will subsequently discuss in the paper.

Problem	Problem Statement
Dominos	A domino is a small, thin rectangular tile that has dots on one of its broad faces. That face is split into two halves, and there can be 0 through 6 dots on each of those halves. Suppose you want to make a set of dominos (i.e., include every possible domino). How many distinguishable dominos would you make for a complete set?
Language Books	You have 4 different Russian books, 5 different French books, and 6 different Spanish books on your desk. In how many ways can you take two of those books with you, if the two books are not in the same language?
CATTLE	How many arrangements of the letters in the word CATTLE have the two T's appearing together either at the beginning or the end of the word?
Quiz Questions	On a quiz, there are 8 multiple choice questions, each of which has 4 possible answers (A, B, C, and D). In how many ways could a student complete this quiz?
Increasing Sequence	You want to make a strictly increasing sequence of length 3, using digits ranging from 0 to 9. How many such sequences are there?

Table 1 – Phase 1 tasks – Initial problem solving

Phase 2 (Sessions 5-7): Reinventing Counting Formulas. Our aim in Phase 2 was to see if the students could reinvent each of the four basic counting formulas. To foster this reinvention, we chose tasks that we believed would motivate generalization, in that the solution sets are too large to enumerate easily via listing. In Table 2, we provide the tasks associated with the reinvention of each formula.

Goal Formula	Problem Statement
$n!$	In the downtown public library, there are 648 books in the children's section. In how many different ways can all of those children's books be arranged on the shelves of the library?
n^r	There are 40 houses in the neighborhood, and they each need to be painted this summer. There are 157 paint colors available. In how many different ways could all of the houses be painted?
$\frac{n!}{(n-r)!}$	There are 19,000 fans at a basketball game. Throughout the game, fifty randomly chosen fans are going to be given fifty different prizes. How many possibilities are there for how the prizes can be distributed?
$\frac{n!}{(n-r)!r!}$	There are 19,000 fans at a basketball game. After the game, fifty fans are going to be chosen randomly to meet the team. In how many ways can these fifty fans be chosen?

Table 2 – Phase 2 tasks – Reinventing four formulas

Phase 3 (Sessions 8-10): Using the Formulas to Solve New Problems. After the students had reinvented the formulas, we wanted to provide opportunities to apply them. Would they use the formulas? If so, would they apply them blindly, or carefully justify their use?

We selected nine tasks, which are included in Table 3. The tasks chosen reflected what we believed to be fairly straightforward instantiations of arrangements, permutations, and combinations. Although students might make unanticipated connections between problems, we nevertheless chose these problems according to which formula/solution type might most naturally fit them (from our perspective), providing at least two problems matching each formula the students had reinvented.

Problem	Problem Statement
SMOTHERING	How many ways are there to rearrange the letters in the word SMOTHERING?
iPhones	In a shipment of 1500 microprocessors, 95 are defective. In how many ways can we select a set of 50 non-defective microprocessors?
License Plates	How many 6-character license plates can be formed using upper-case letters and the digits 0-9?
Bits	Consider binary strings that are 256 bits long. How many 256-bit strings contain exactly 75 0's?
Horses	20 horses are running in the Kentucky Derby. How many options are there for which horses could finish with Win, Place, or Show?
Coin Flip	A fair coin is flipped 36 times. How many outcomes are possible? How many outcomes have as many heads as tails? How many outcomes have a head on the fifth toss?
Lollipops	There are 50 children, and there are 10 identical lollipops to give to the children. How many ways could the lollipops be distributed if no child can have more than one lollipop?
Kickball	There are 30 kids who want to play kickball in gym class, and in a game of kickball there are 9 positions on the field at a time. How many ways are there for 9 of these kids to play in a kickball game?
Paintings	In the Portland Art Museum, there are 25,000 paintings. In how many ways can these pieces of art be displayed on the walls of the art museum?

Table 3 – Phase 3 – New problems

Analysis. As the experiment proceeded, we conducted an *ongoing analysis* that included reviewing the videotape of each session and constructing a *content log* of that session. In creating these content logs, we paid particular attention to students' articulated thoughts that seemed to provide them with leverage, the voicing of concerns or perceived hurdles that needed to be overcome, and signs of/causes for progress. Our ongoing analysis informed our decisions about tasks for subsequent sessions. We also conducted a *retrospective analysis* (Cobb, 2000) in which we reviewed the entire corpus of data at a deeper level, so as to refine our descriptions of thematic elements present in the students' reasoning. This retrospective analysis included re-watching all of the videos, enhancing the content log for each session, and noting recurring phenomena and themes. Once we had identified a handful of themes, we reviewed and transcribed key portions of the video in order to investigate aspects of the data related to each theme.

Results

We divide the results into two findings: First, students' use of sets of outcomes supported their successful reinvention of the four formulas and increased rate of success on subsequent problems, and, second, overreliance on sets of outcomes and patterning may preclude the development of the MP. In describing both findings, we present the students' work on one or two problems that is representative of their overall work.

Finding 1: Students' use of sets of outcomes supported their successful reinvention of the four formulas and markedly increased rate of success on subsequent problems

Phase 1 (Sessions 1-4): Solving Initial Counting Problems. The students' work on this phase was marked by a focus on writing outcomes, and, in some cases, relying on patterns to determine the answer to the problem. The Domino problem was their first exposure to solving counting problems together, and we detail their work to highlight a couple of important norms that were established. Their work on this problem is representative of their work with outcomes throughout the experiment.

Thomas began by writing out a row of seven dominos, writing a 0 over a 0 to represent the 0:0 domino (Figure 4). They then had the following exchange, which set the stage for prioritizing listing over a search for an equation.

Robin: "Isn't there an equation for that?"

Thomas: "I bet there is."

Robin: "I don't remember it…Isn't it something like, what are those things called? … Like combination equations…How should we do this?"

Thomas: "I don't know, I don't know the equation, so…(trails off)."

After this exchange, they did not pick up on this search for an equation, and the idea of simply looking for an equation was tacitly set aside. Without knowing what such a formula might look like, they instead proceeded with listing more outcomes. Thomas said, "For each of the top half, there would be seven that match with it. So we'd have to do one for every one of them, 7, 7, 7, 7, 7, 7, 7," suggesting repeated addition. They then started to write a second row of seven dominos (1:0 through 1:6), but they immediately realized that they had already counted 0:1. Thomas suggested that there thus would not be 49 dominos, but he had not yet recognized a correct pattern, saying, "so this one would have seven and then each of the rest would have like 6." At this point, Robin suggested that they write out more dominos, which they did.

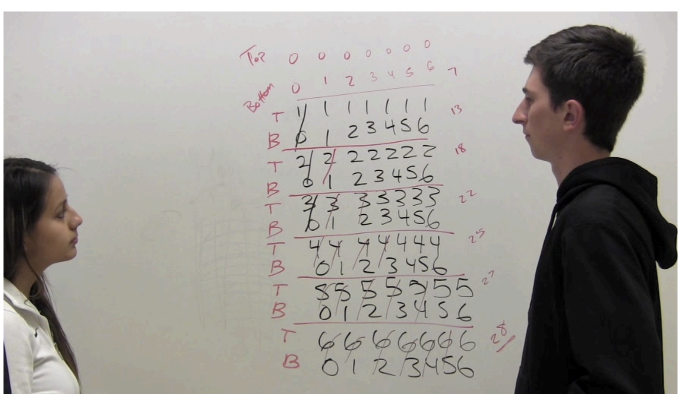


Figure 4 – Thomas and Robin's work on the Dominos problem

As they did, they listed the second row of dominos with 1's and crossed off the duplicate 1:0 domino. They then began to write out the dominos with 2's, and they crossed out 2:0, but as they wrote 2:1 Thomas said, "Ooooh, one more cancels each time." This suggested that he had not initially realized the pattern until he was actually writing out that particular outcome, and that the listing helped him recognize the duplication. They then wrote the entire list (Figure 4) and crossed out all of the duplicates. They recognized that "one more cancels each time" and ultimately summed to find the correct total 7+6+5+4+3+2+1 = 28.

This problem highlights the value of articulating and listing outcomes, particularly in recognizing issues involving overcounting. Indeed, it seems as though Thomas might not have recognized that "one more cancels each time" if he had not listed more outcomes. Additionally, we suggest that this problem (and the use of outcomes) was key in determining what they took as an acceptable answer to a counting problem. They were confident in their answer of 28 because they had listed all of the outcomes, and this work set the stage for much of Phase 1. By the time they finished this phase, they had solved all 10 problems correctly.

Phase 2: Reinventing Counting Formulas. To exemplify their work in this section, we very briefly discuss their reinvention of the formula for the number of arrangements of r objects from n distinct objects: $\frac{n!}{(n-r)!}$. Posed with the third task in Table 2, the students noted that they wanted to multiply 19,000 by 18,999 by 18,998, and so on, all the way down to 18,951. They eventually recognized that they could more efficiently write that product as a quotient of factorials. This led them to conjecture that their answer would be $\frac{19000!}{18950!}$. To check their conjecture, they tried some smaller examples, which gave them growing confidence that their conjecture was sound. This led them to generalize their conjectured formula to the one seen on the right-hand side of Figure 5.

By the end of Phase 2, the students had successfully reinvented each formula, using their own notation. Table 4 shows the formulas that the students came up with, as compared to standard textbook formulas. Despite the somewhat idiosyncratic appearance, the students' notation was meaningfully connected to their experience; for instance, the f in the last two formulas stood for fans at a Blazer game, which was the context in which the expressions were developed.

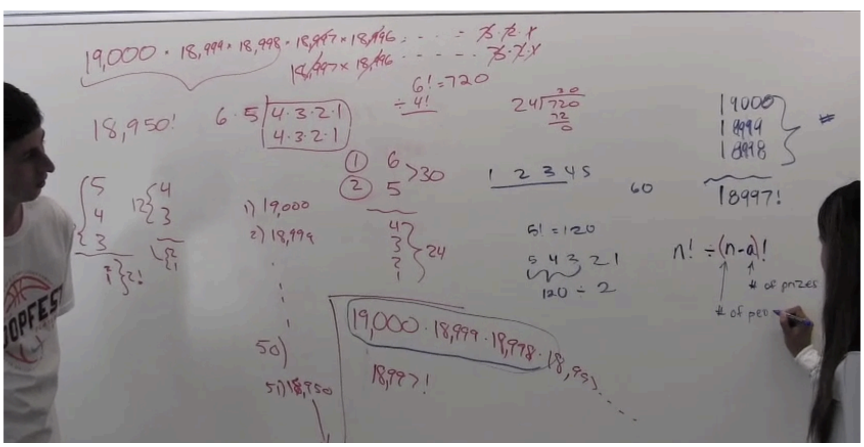


Figure 5 – Reinventing permutations of r objects from n objects

Textbook Formulas	$n!$	n^r	$\frac{n!}{(n-r)!}$	$\frac{n!}{(n-r)!r!}$
Students' Formulas	$n!$	a^b	$\frac{n!}{(n-f)!}$	$\frac{n! \div (n-f)!}{f!}$

Table 4 – Textbook formulas and students' reinvented formulas

Phase 3: Using the Formulas to Solve New Problems. What is noteworthy in this phase is that the students still used outcomes to determine which formula to use. Even when armed with their new formulas, the students did not blindly apply them. Instead, they first articulated what an outcome was, discussed the nature of that outcome (whether repetition was allowed, whether order was relevant), and used that discussion to determine which formula to use. In this way, they solved eight of the nine problems correctly. Figure 6 shows their (correct) work on the Kickball problem, showing how they first wrote down an outcome and decided that different orderings of the kids yielded different outcomes. This suggested permutations rather than combinations, again demonstrating how their facility with outcomes positively affected their work, helping to decide which formula to apply when.

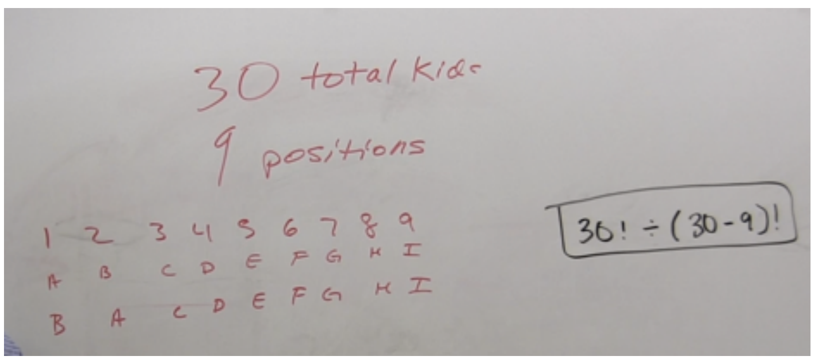


Figure 6 – The students' work on the Kickball problem

Summary of Finding 1: To elucidate our first finding, we want to emphasize three aspects of their usage of sets of outcomes. First, throughout their work on these problems, the students used sets of outcomes frequently and with great effect. As discussed in their work on the Domino problem, the students reasoned about outcomes, to a much greater extent than previous studies have suggested (Lockwood, 2011, 2013). Second, the students did use such reasoning to successfully reinvent the formulas, and, even more, they went on to solve problems successfully using these formulas. That is, they did go through the tasks and were able to arrive at formulas to which they had not previously been exposed. Again, we find this to be a notable finding, demonstrating both that students are capable of such activity, and that they were able to use those formulas in subsequent counting activity. Finally, students were remarkably successful at solving these problems. Indeed, through the course of the teaching experiment, they solved 29/30 (96.7%) problems correctly. We cannot overstate how noteworthy and impressive these results are, given low success rates and difficulties that have frequently been reported in other studies.

Finding 2: Overreliance on sets of outcomes and patterning may preclude the development of the MP

Given the students' success described above, it is reasonable to expect their work to be characterized by rich and frequent justification via the multiplication principle. However, our analysis indicates that the students' work was surprisingly not based on the multiplication principle, but instead was almost entirely based on empirical patterning. These findings reveal unexpected phenomena that suggest new insight into Lockwood's (2013) model. This point can perhaps best be illustrated by presenting the students' work on the Quiz Questions problem.

Quiz Questions. As discussed above, a natural approach to solving the Quiz Questions problem is to use the multiplication principle; we can argue that there are eight independent stages to the problem, each of which have four possibilities, and so the answer is 4^8. Indeed, in our prior experience giving this problem to students, it is natural for them to use this positional reasoning and to use multiplication in this way to solve the problem, and we expected that Thomas and Robin might reason similarly.

Instead, however, Thomas and Robin created a table with 8 columns for the questions 1 through 8, and they tried to enumerate the possibilities by keeping everything constant and then changing the last digit (Figure 7). Accordingly, they wrote rows of AAAAAAAA, AAAAAAAB, AAAAAAAC, and AAAAAAAD, and found there to be four options for changing the last letter. Even though they had the eight different questions labeled, at no point did they talk about the number of choices they had for each position, and they did not appear to think of constructing an outcome in stages by considering that there were four choices for each question. The excerpt below, from the beginning of Session 2 (in Phase 1), shows how they arrived at the correct answer to the problem; but what is noteworthy is that they never seemed to reason using the multiplication principle. Indeed, as we will see, they were driven to the correct answer of 4^8 by noticing how quickly things were growing (suggesting exponentiation) and verifying this guess by noticing a pattern.

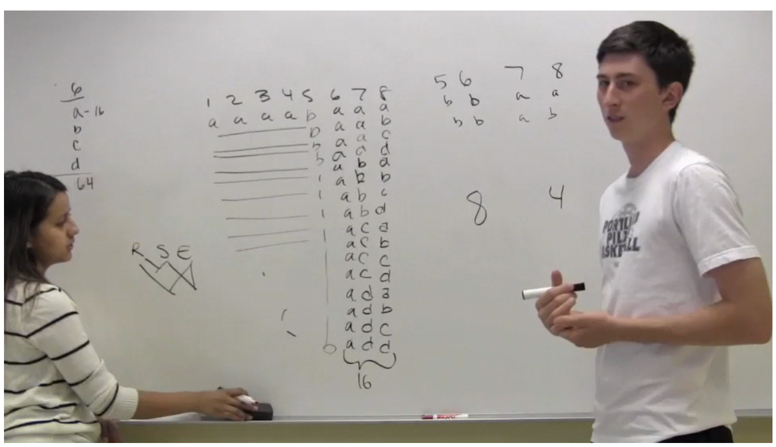


Figure 7 – The students' work on the Quiz Questions problem

After much work, Thomas had written an 8 and a 4 on the board, and he articulated his frustration about the magnitude of the answer. His work here suggests that Thomas was motivated to pursue exponentiation because he realized the options were growing quickly.

Thomas: "We're getting really frustrated trying to write out all of the possibilities, because we're just noticing it's just going to keep growing and growing. So we're trying to think of a way we can just either multiply them [referring to the numbers 4 and 8], or do something with them. I wonder, because it's growing really fast…I wonder if you would [*writes an exponent of 4 on the 8 - they both laugh*]. It's just, they're growing really fast, and so maybe exponentially, 'cause then that would give us a really large number."

After a bit more thought, and in describing the growth of his pattern, Thomas adjusted the conjecture and wrote 4^8.

Thomas: "Well because now, maybe it's not the 8 to a power, because we see that there's 4, changing just this last column, 'cause it's each one of the possibilities. And then when we move over a row, then we get like a total of 16 possibilities, which is 4 times 4, or 4 squared, and so maybe if we go out, should we do this one? Oh we got 64, which is 4 cubed, and so gosh."

Robin: "What if it's 4 to the *n*?"

Thomas: "We've changed 3 questions and we've got 4 cubed. So if we changed the 4th question we'd hope to get 4 to the 4th. Ooh, my goodness [*writes and calculates* $64 \times 4=256$]."

Robin: "Yeah so what if it's just 4 to the *n*, I could see that."

Thomas: "It's going to be hard to check it, because already we didn't want to count up like this fourth column, which we think would give us 256, it's just going to be hard to check if we're right or not. It fits these first three. Um, I don't know."

While they ultimately decided that the answer should be 4^8, they went on to say that they did not see how else they could check their answer, noting "we can only go up to the third row without missing anything." There was not a sense that they could argue what was happening in terms of why this answer might make sense more generally, aside from the pattern they had detected. We thus see in this problem that while they arrived at the correct answer of 4^8, this answer was entirely based on a pattern they had empirically established through writing outcomes. They had found four possibilities by listing outcomes, and then they found 16 by listing outcomes, and so that, combined with the focus on exponentiation, suggested to them that the answer was 4 to a power. It is not the case that they reasoned about the number of options possible for each question, nor did they really seem to have ownership of why the solution made sense.

This absence of MP reasoning is representative on their work on almost all of the problems that involved large numbers, throughout the entire experiment. In fact, their reinvention of the formulas and their subsequent work on new tasks all suggest a similar, consistent method. The ability to match patterns and reason from sample outcomes was reliable for them, and their work on every subsequent problem continued to provide repeated and consistent evidence for the fact that they did not use the multiplication principle in their reasoning. Accordingly, although they almost invariably eventually arrived at the correct answer by way of applying a correct formula (and using sets of outcomes to do so), their understanding of why the answer made sense was perhaps incomplete.

In summarizing the results of the experiment, then, there is a tension between two different aspects of the overall narrative that we can describe. On the one hand, these students were very successful. Given the rates at which students typically correctly solve counting problems, even in studies in which they have ample time to work on a problem, a 96.7% success rate is very impressive. Additionally, their ability to reinvent formulas, and their impressive facility with sets of outcomes are laudable. We want to celebrate these students' accomplishment in this regard and to present our ideas for why they were so successful. On the other hand, we do not feel that their high success rates tell the entire story, as there were also serious limitations to their approach. Specifically, because of their strong connection to the set of outcomes, and because this often enabled them to rely heavily on patterns, they did not encounter a need to develop strong multiplicative thinking to explain what they were doing. As a result, they often lacked the kind of robust reasoning and justification that would be expected of students who had truly understood the material typically presented, say, in an introductory course in discrete math.

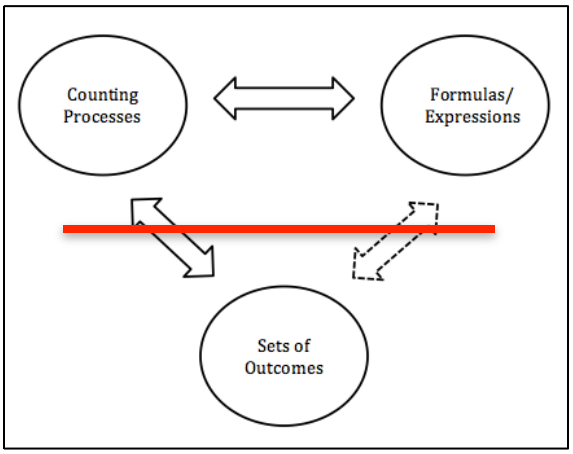


Figure 7a

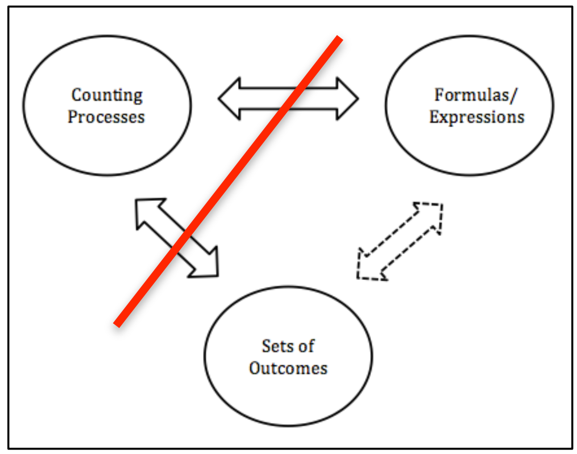


Figure 7b

The students' work also shed some new light on Lockwood's (2013) model. In particular, prior to this study, students' struggles with counting problems might be frequently characterized by Figure 7a, in which students commonly neglect sets of outcomes and instead formula-match or rely on memorized problem features. However, the students in our study were not making these same kinds of mistakes. Instead, in terms of the model, we would characterize their work by the diagram in Figure 7b. Our students did not display a grasp of the multiplication principle (in fact, their work suggests they did not have the MP as a tool), and they did not use a counting process like the MP to justify their work. Instead, they relied entirely on empirical patterning (which may be similar to Harel's (2001) *result pattern generalization*). Indeed, as mentioned above, prior to this study, the relationship between sets of outcomes and formulas/expressions was not well understood. Our findings thus make a theoretical contribution to the model, suggesting that *patterning* is an appropriate description for how to characterize the relationship between sets of outcomes and formulas/expressions. Moreover, the students in this study provide an existence proof that, given appropriate circumstances, this can develop in students as a practical form of reasoning about counting.

Discussion and Next Steps

A number of studies (e.g., Lockwood, 2011, 2013) have already shown that sets of outcomes are an important feature of combinatorial reasoning and can support good counting practices, and we still claim that this is true. As we have indicated, the students productively reasoned about outcomes in almost every counting situation in which they found themselves – in formulating the answer to a problem in Phase 1, in developing patterns that contributed to generalization in Phase

2, and in determining which formula best fit their current situation in Phase 3. Looking at and arguing about outcomes was a fundamental aspect of their deciding which formulas might be appropriate, and this allowed them to avoid nonsensical answers and led them away from the temptation simply to apply the formulas blindly. The study thus provides strong evidence for how useful focusing on outcomes can be in counting, particularly for novice counters

However, this study also reveals that sets of outcomes are not sufficient in helping students make sense of and justify their counting activity, and we highlight two potential issues with these students' attention to outcomes in absence of counting processes like the MP. First, these students showed that patterning may not lead to justification. If indeed we seek to have students who deeply understand their counting activity (which is a clear goal for developing proficient counters), we need to emphasize more than just patterning. In addition, our study suggests that an overreliance on outcomes may preclude the development and use of the MP, the understanding of which, as we have mentioned, is central to the entire subject. Therefore, while work with outcomes can be useful for students, caution should also be taken so students do not rely on outcomes to the point that they ignore deeper and more conceptual aspects of their work.

There are a number of natural next steps in our program. We need to investigate the ways students connect their counting processes with their sets of outcomes, and we feel that the multiplication principle is a key aspect of this connection. Therefore, we want to investigate the principle more deeply, studying students' conceptualizations and development of it as a tool for counting. This may involve reinventing the multiplication principle specifically, which we believe may be more productive for students than having them simply reason about statements they are simply given or told.

Ultimately, we want to move toward designed-based research that targets the development of instructional tasks and sequences that can help students be more conceptually grounded in their counting activity. Such research would follow the trajectories of researchers in other undergraduate content areas, such as linear algebra, differential equations (e.g., Rasmussen & King, 2000), and abstract algebra (e.g., Larsen, 2013). This study has laid the groundwork for such subsequent investigations in which we can explicitly target the development of particular concepts and ideas (such as the multiplication principle, or sets of outcomes) that we feel might be important aspects of students' combinatorial thinking.

References

Annin, S. A., & Lai, K. S. (2010). Common errors in counting problems. *Mathematics Teacher, 103*(6), 402-409.

Batanero, C., Navarro-Pelayo, V., & Godino, J. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*, 181-199.

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In R. A. Lesh & A.E. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 307–333). Hillsdale, NJ: Erlbaum.

Eizenberg, M. M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning, 6*(1), 15-36.

Fischbein, E., & Gazit, A. (1988). The combinatorial solving capacity in children and adolescents. *ZDM, 5*, 193-198.

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht, The Netherlands: Reidel.

Gravemeijer, K. (1998). Developmental research as a research method. In A. Sierpinska & J. Kilpatrick (Eds.), *Mathematics Education as a Research Domain: A Search for Identity* (pp. 277-296). Dordrecht, The Netherlands: Kluwer.
Gravemeijer, K., Cobb, P., Bowers, J., and Whitenack, J. (2000). Symbolizing, modeling and instructional design. In P. Cobb, E. Yackel and K. McClain (Eds.), *Symbolizing and Communicating in Mathematics Classrooms*, Erlbaum, Mahwah, NJ, pp. 225–273.
Hadar, N., & Hadass, R. (1981). The road to solve combinatorial problems is strewn with pitfalls. *Educational Studies in Mathematics, 12,* 435-443.
Harel, G. (2001). The Development of Mathematical Induction as a Proof Scheme: A Model for DNR-Based Instruction. In S. Campbell & R. Zaskis (Eds.). Learning and Teaching Number Theory, Journal of Mathematical Behavior. New Jersey, Ablex Publishing Corporation (pp. 185-212).
Larsen, S. (2013). A local instructional theory for the guided reinvention of the group and isomorphism concepts. *Journal of Mathematical Behavior*. Doi: 10.1016/j.jmathb.2013.04.006 .
Lockwood, E. (2011). *Student Approaches to Combinatorial Enumeration: The Role of Set-Oriented Thinking.* Unpublished doctoral dissertation, Portland State University.
Lockwood, E. (2013). A model of students' combinatorial thinking. *Journal of Mathematical Behavior, 32,* 251-265. Doi: 10.1016/j.jmathb.2013.02.008.
Maher, C. A., Powell, A. B., & Uptegrove, E. B. (Eds.). (2011). *Combinatorics and Reasoning: Representing, Justifying, and Building Isomorphisms.* New York: Springer.
Martin, G. E. (2001). *The Art of Enumerative Combinatorics*. New York: Springer.
Oehrtman, M., Swinyard, C., & Martin, J. (2014). Problems and solutions in students' reinvention of a definition for sequence convergence. *Journal of Mathematical Behavior, 33,* 131-148.
Rasmussen, C., & King, K. (2000). Locating starting points in differential equations: A realistic mathematics approach. *International Journal of Mathematical Education in Science and Technology, 31*, 161-172.
Richmond, B. & Richmond, T. (2009). *A Discrete Transition to Advanced Mathematics.* Providence, RI: American Mathematical Society.
Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates.
Swinyard, C. (2011). Reinventing the formal definition of limit: The case of Amy and Mike. *Journal of Mathematical Behavior, 30,* 93-114.
Tall, D & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151–169.
Tucker, A. (2002). *Applied Combinatorics* (4th ed.). New York: John Wiley & Sons.

THE CONSTRUCTION OF COHOMOLOGY AS OBJECTIFIED ACTION

Anderson Norton
Virginia Tech

The purpose of this paper is to investigate a theory about the nature of mathematical development, in which mathematics is characterized as the objectification of action. Informed by existing research on how students construct new mathematical objects, we consider as an example the psychological construction of cohomology and related objects of algebraic topology. This example extends neo-Piagetian theories of mathematical development from elementary school to graduate-level mathematics, while integrating existing research on students' learning of abstract algebra. Results of the investigation affirm the objectification of action as a distinguishing feature of mathematics in general, while indicating the kinds of mental actions that undergird the objects of advanced mathematics.

Key Words: Abstract Algebra, APOS Theory, Constructivism, Reflective Abstraction, Reification

"Mathematics is the science of actions without objects, and for that, of objects we can define through action." Paul Valéry (1973, p. 811).

When fields' medalist William Thurston endeavored to address the plight of mathematics education in the United States, he shared the following personal anecdote:

> I remember as a child, in fifth grade, coming to the amazing (to me) realization that the answer to 134 divided by 29 is 134/29 (and so forth). What a tremendous labor-saving device! To me, '134 divided by 29' meant a certain tedious chore, while 134/29 was an object with no implicit work. I went excitedly to my father to explain my major discovery. He told me that of course this is so, a/b and a divided by b are just synonyms. To him it was just a small variation in notation. (Thurston, 1990, p. 5)

Thurston used the story to illustrate the challenge we face, as teachers, when we attempt to unpack the mathematical objects we have constructed. Mathematics education researchers have taken pains to unpack the object of Thurston's example in particular, demonstrating how students begin to understand fractions (and especially improper fractions, like 134/29) as "numbers in their own right" (Hackenberg, 2007). The key to this and similar work has been to identify the mental actions that comprise those objects, thus equipping teachers and researchers with models for how students might construct those objects through activity.

Few students in the United States accomplish what Bill Thurston did (Norton & Wilkins, 2012). In fact, it's possible that Thurston's father did not appreciate his son's revelation because, for him, the fraction 134/29 symbolized nothing more than the division of two whole numbers. On the other hand, if the elder Thurston *had* constructed 134/29 as a number, it's probable that he would have forgotten the labor of that construction, which involves coordinating mental actions of partitioning and iterating within a three-level structure: 134/29 as a unit resulting from 134 iterations of a 1/29 unit, which results from partitioning a whole unit into 29 parts (Hackenberg, 2007). Figure 1 illustrates such a structure for the simpler fraction, 8/3. This structure supports a conception of the improper fraction as an object defined through its size relation with the whole: 8/3 as a number that is eight times as big at 1/3, which has a 1-to-3 size relation with the whole.

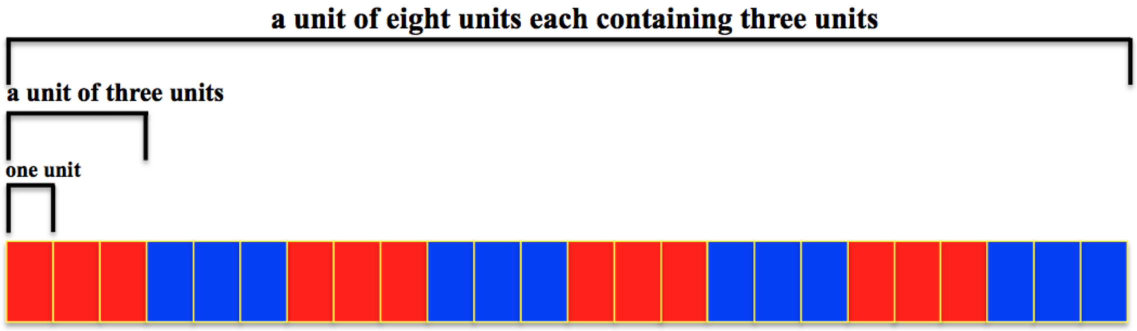


Figure 1. 8/3 as a unit of units of units.

Steffe and Olive (2010) have described this way of conceptualizing improper fractions as an *iterative fraction scheme* (IFS). Whereas we have fine-grained models for describing, explaining, and predicting the construction of improper fractions, few models of this kind exist for advanced mathematics. The scarcity of such models likely owes to two factors: (1) mapping the psychological construction of mathematics requires intensive and longitudinal studies of students' development—studies that, so far, have followed a trajectory from infancy to middle school mathematics; and (2) although schemes seem adequate for building models of development up to that point, modeling students' constructions of advanced mathematics likely requires more complex structures. Here, we will examine construction in an extreme case—cohomology—to identify key mental actions, even if we cannot model the complexity of their coordination.

Theoretical Framework

Inherent in Piaget's genetic epistemology is the idea that mathematical objects arise through the coordination of actions: "The meaning of objects has two aspects: It is 'what can be done with them' either physically or mentally… The meaning of object is also 'what it is made of,' or how it is composed. Here again, objects are subordinate to actions." (Piaget & Garcia, 1986, pp. 65-66). As Tall and colleagues (2000) have noted, several theoretical frameworks for teaching and learning have arisen from this idea, including APOS theory (Dubinsky, 1991), reification (Sfard, 1991), and scheme theory (von Glasersfeld, 1995). Here, we present a broader theoretical framework that builds on such work while aligning more closely with Piaget's characterizations of actions and objects, as well as his characterization of mathematics itself.

APOS Theory

Dubinsky and colleagues (e.g., Dubinsky & Lewin, 1986) developed APOS theory as a means of applying Piaget's constructivist epistemology to research on undergraduate mathematics education. In particular, they demonstrate how mathematical actions may become *reflectively abstracted* as advanced mathematical objects and schemas. Their central tenet is that "mathematical knowledge consists in an individual's tendency to deal with perceived mathematical problem situations by constructing mental *actions*, *processes*, and *objects* and organizing them into *schemas* to make sense of the situations and solve the problems" (Dubinsky & McDonald, 2001, p. 2). In this framework, *actions* are defined as transformations of tangible objects (including diagrams and written symbols) and might include carrying out the steps of an algorithm, such as computing the left cosets of a particular algebraic group. Reflecting on such actions allows the individual to internalize them as mental *processes* that the individual can imagine performing, without the need for tangible objects. Similar to Piaget (1970b), Dubinsky and McDonald (2001) argue that this internalization allows students to reverse and compose actions. The process becomes an *object* for an individual when he or she can symbolize it and purposefully act upon it. “Finally, a schema for a particular mathematical concept is an individual’s collection of actions, processes, objects, and other schemas which are linked by some general principles to form a framework in the individual’s mind” (p. 3).

Reification

Following Dubinksy (1986), Sfard (1992) further elaborated on Piaget’s (1970a) notion of reflective abstraction by prescribing three stages through which students progress from engaging in mathematical processes to producing mathematical objects. To illustrate, Sfard provided an extended example from the historical development of number: from natural numbers, to positive rational numbers, to positive real numbers, to real numbers, and finally to complex numbers. She argues that each step-wise development has depended upon stages of *interiorization*, *condensation*, and *reification*. In particular, in the production of rational numbers, processes involving the division of natural numbers become *interiorized* so that they “can be carried out in mental representation” (p. 18, from Piaget, 1970a). Then they are *condensed* so that they can be combined with other processes, such as measurement. Finally, they are *reified*, or objectified, as a static structure on which to perform further processes, as in the development of positive real numbers. In fact, we can find evidence of this kind of development in the personal experience shared by Thurston: Whereas 129/34 had been a laborious process to perform, perhaps interiorized and condensed over a period of learning, in an instant it became reified as an object or “compact whole” (Sfard, 1992, p. 14). Unfortunately, Bill Thurston’s father did not appreciate this “quantum leap” (p. 20) from process to object, which we might explain in either of two ways, as discussed later in this section.

Scheme Theory

Sfard did not make use of Dubinsky’s *action-process* distinction, allowing processes to include actions, whether carried out physically or mentally. Neither did she make use of *schemas*. In contrast, scheme theory relies on a different characterization of action and utilizes a construct similar to Dubinsky’s schema, but does not explicitly address the production of *objects*. von Glasersfeld (1995) described a *scheme* as a three-part structure: an assimilatory template of situations that might activate the scheme, a coordinated collection of mental actions carried out by the scheme, and an expected result from acting in the situation. Although Dubinsky’s and Sfard’s frameworks would include such actions, von Glasersfeld’s description of mental action drew more heavily and narrowly from Piaget. For example, in contrast to the more formal mathematical actions of dividing and measuring described in Sfard’s analysis of how students construct positive rational numbers, a scheme theoretic perspective would focus on the *psychological* actions that undergird them.

Actions and Objects

In an attempt to characterize the nature of mathematical objects and their construction, Tall and colleagues (2000) reviewed each of the frameworks described here and, noting the common theme of encapsulated actions, sought to describe how actions become objectified. Here, we broaden these frameworks and extend their purpose by arguing that mathematics *is* the objectification of action—this is what makes our field unique and, in some sense, infallible.

Unlike other sciences, languages, or any other field of study, all of the objects of mathematics are based on actions and their coordination so that, ultimately, mathematical claims are about nothing but the mental actions we can perform. If these actions correspond to (or even predict) experiential reality, it is only because we, as humans, have evolved to operate within the world we experience (Piaget, 1971/1970).

Piaget's epistemological research draws a fundamental distinction between two kinds of thought: figurative and operative. Whereas figurative thought pertains to empirical abstractions of "perception, imitation, and mental imagery" (1970a, p. 14), operative thought is the domain of mathematics. It pertains to reflective abstractions of one's coordinated activity in the construction of mental actions and structures. Unlike figurative objects (such as colors and drawings), operative objects remain dynamic on the basis of the actions that comprise them and the structures that organize them. Moreover, constructing such objects opens new possibilities for action, so that mathematics continually builds upon itself in alternating layers of actions and objects. Figure 2 illustrates the basic character of operative thought.

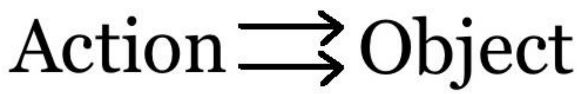


Figure 2. Mathematics as objectified action.

The top arrow in Figure 2 indicates that actions become reflectively abstracted as objects. The bottom arrow indicates that, as objects, these objectified actions can be acted upon. This pattern lies at the heart of Piaget's epistemology of mathematics and can also be found Sfard's reification and Dubinsky's APOS theory. What Sfard and Dubinsky do not address is how interiorized actions become organized within *psychological* (rather than formal mathematical) structures—the subject of Piaget's structuralism.

Structuralism

Structuralism focuses solely on operative thought, as an attempt to explain how children develop logico-mathematical reasoning. In addition to schemes (discussed above), Piaget (1970b) posited algebraic group-like structures that organize mental actions into reversible and composable systems. For example, students who have constructed mental actions of partitioning and iterating might organize them as inverse elements within a "splitting group", where iterating a part five times undoes the mental action of partitioning a continuous whole into five parts (Norton & Wilkins, 2012). They might also engage in recursive partitioning, in which partitioning is both an action and the object of that action (e.g., partitioning a continuous whole into three parts and then partitioning each of those parts into five parts to produce fifteen parts in the whole). Recent research (ibid) indicates that this group-like structure is necessary for the construction of IFS—the way of operating Thurston apparently constructed in fifth grade.

Although Piaget's epistemology (including his structuralism) equates logico-mathematical thought with operative thought, much of what happens in mathematics classroom involves figurative thought as well (Thompson, 1985). When the link is broken between a student's mental actions and the objects of a mathematical lesson, the student has little recourse but to engage in figurative thought. Sfard and Linchevski (1994) referred to this kind of engagement as the *pseudostructuralist approach*: "The new knowledge remains detached from its operational underpinnings and from previously developed systems of concepts" (p. 221). Moreover, Thompson (1985) has argued that students foreground some objects of mathematical discussion as operative—acting on them and deconstructing them into their constituent actions—while placing other objects in the background, as figurative. For example, functions might be operative in the context of high school algebra, as students act on covarying quantities and attempt to establish them as invariant relationships, but functions might be treated as figurative within cohomology, where they are elements of a group. In any case, what constitutes operative thought depends upon the available mental actions of the individual and her goals within the activity. Thus, we can say the same for mathematics.

Research on Abstract Algebra from an Action-Object Perspective

Action-object perspectives (especially Sfard's reification and Dubinsky's APOS theory) have gained strong influence in research on undergraduate mathematics education (RUME). Here, we review RUME studies from an action-object perspective that focus on concepts related to abstract algebra, and therefore related to algebraic topology and cohomology (for which no direct mathematics education research exists).

In a study on how college mathematics majors learn group isomorphism, Leron, Hazzan, and Zazkis (1995) drew a distinction between students who understood "the relation of two groups being isomorphic" and those who understood "the object of isomorphism" (p. 154). They identified three phases in students' transition from the former, action/process conception, to the latter, object conception: (1) concepts that reference the student doing something; (2) concepts that reference a process that could be carried out by anyone; (3) concepts that make claims of subject-independent existence. As students struggled to progress toward an object conception of isomorphism, the researchers noticed them "craving for canonical procedures and their fear of loose or uncertain procedures, indeed, procedures with any degree of freedom" (p. 171).

In a similar study with high school teachers, Dubinsky, Dautermann, Leron, and Zazkis (1994) focused on the interconnected layers of objects within group theory—group, subgroup, coset, normality, and quotient group—and their dependency on existing concepts of set and function. The teachers tended to begin by treating groups as sets on which to act and only later considered the role of a binary operator (function) in defining groups as objects. In line with Leron, Hazzan, and Zazkis (1995), the researchers noted the need for a concept of isomorphism in order to construct "group as an equivalence class of isomorphic pairs [of sets and functions]" (Dubinsky et al., 1994, p. 290). They also found that teachers construct subgroups in parallel with groups, as functions with a restricted domain. However, the teachers were generally not successful in constructing quotient groups, which the researchers attribute to difficulty in objectifying the process of forming cosets—a prerequisite construction for treating cosets as elements of a group. This difficulty was associated with teachers' tendency to conflate normality and commutativity.

Hazzan (1999) found that undergraduate students deal with the complexity of abstract algebra by "reducing the level of abstraction" (p. 71). Students do this in three distinct ways: (1) by basing arguments on more familiar mathematical entities (such as sets, rather than groups); (2) by dealing with single elements within a more complex collection (for example, working with a representative element within a quotient group, rather than the quotient group itself); and (3) by reducing objects to the actions that comprise them. Although the three methods are closely related, the third method aligns most directly with an action-object perspective. In line with the study by Leron, Hazzan, and Zazkis (1995), students can reduce the complexity of an entity by imagining actions they can perform to build it up. For example, one student dealt with quotient groups, G/H, by referencing the imagined activity of taking all elements of the normal subgroup, H, and choosing an element from the group G by which to multiply them on the right.

Other studies have demonstrated the efficacy of an action-object perspective as a pedagogical tool (e.g., Asiala, Dubinsky, Mathews, Morics, & Oktac, 1997; Brown, DeVries, Dubinsky, & Thomas, 1997). For example, Asiala and colleagues (1997) reported on the effectiveness of an abstract algebra course that explicitly attended to students' progressive constructions of actions, processes, objects, and schema. In particular, they described an action conception of coset as one in which students could work with simple and familiar groups/subgroups to build the coset. Students progress to process conceptions of coset when they can imagine computing the products just as the student in the example provided above, from Hazzan (1999). Students can then progress to object conceptions, in which they do not need to focus on the actions of building the coset and instead act on the coset itself. Finally, a coset schema is formed as a network of actions, processes, objects, and schemas, by relating cosets to concepts of groups, subgroups, normality, and quotient groups.

The Construction of Cohomology

When considering the complexities of an advanced mathematical idea, diagrams can provide some indication of their organization. Specifically, Figure 3 represents various components of cohomology and their relationships. However, for most of us, these components and relationships remain figurative rather than operative because they do not symbolize mental actions that we perform, nor objects that we act upon. The situation is completely analogous to that faced by middle school students as they begin engaging in algebraic manipulation without reference to underlying mental actions. For example, students commonly solve equations of the form *ax*=*b* by subtracting *a* from both sides of the equation. Correcting students' behavior in these instances is unproductive in terms of supporting algebraic reasoning. We need to address the source of the problem, that algebraic manipulations should become a proxy for underlying mental actions on previously constructed objects.

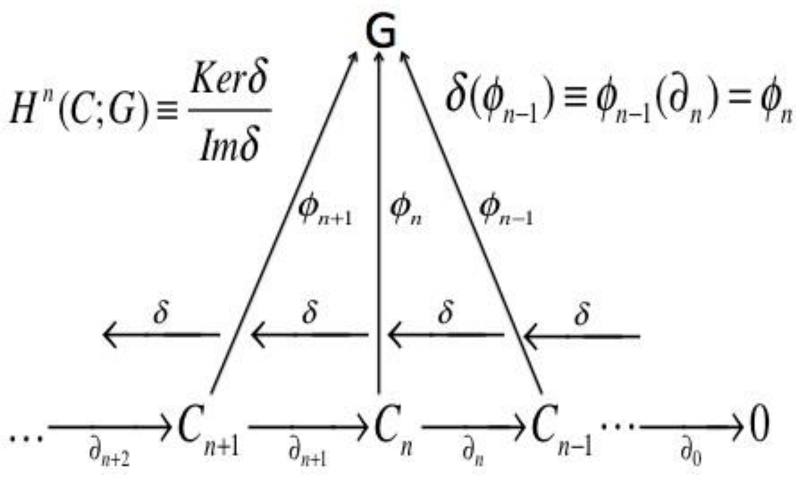


Figure 3. Diagram of cohomology

Previous research has suggested that constructing concepts in abstract algebra relies on having constructed functions and sets as objects first (Dubinsky, Dautermann, Leron, & Zazkis, 1994). Students tend to begin by treating groups as sets on which to act and only later consider the role of a binary operator (function) in defining groups as objects. Also, researchers have noted the interdependency of groups and isomorphisms in constructing “group as an equivalence class of isomorphic pairs [of sets and functions]” (Dubinsky et al., 1994, p. 290). Figure 3 begins at this stage, where C_n, represents a free abelian group generated by the set of n-dimensional triangles (e.g., vertices, edges, triangles, tetrahedras, etc.) used to build up the topological space under consideration. $\P_n$ represents a “boundary map” from C_n to C_{n-1}: a homomorphism that maps each n-dimensional triangle to its boundary (e.g., the boundary of an edge is the difference between its vertices, v_2-v_1). G represents another, selected group, and the various ϕs represent functions from C_n to G. Suppose these are objects for us, in the sense that Asiala and colleagues have described (1997): We can act on them and unpack them to their constituent actions (as opposed to figurative objects on which we might act but are not themselves composed of actions). Now consider the chain complex—the abelian groups, C_n, and the boundary maps, $\P_n$, between them—as an algebraic procedure. Thus, Figure 3 serves to identify the boundary between algebraic objects and actions, even though we have not yet identified what psychological actions might undergird procedures associated the chain complex.

The Circle

To proceed, we might compute the homologies of familiar spaces. Computing homology allows us to focus on objectifying the chain complex while reducing further complexity introduced by *co*homology: the inclusion of the “ϕ” functions to group G and the coboundary maps, δ. Let us begin by computing the homology of the circle. This decision can be interpreted as an attempt to “reduce the level of abstraction” by dealing with a familiar entity (Hazzan, 1999), which might also make it easier to geometrically interpret the results of our algebraic computations. In particular, it is easy to see how a circle can be continuously deformed into a triangle, with three vertices and three edges. Thus, the chain complex becomes $0 \rightarrow <e_1, e_2, e_3> \rightarrow <v_1, v_2, v_3> \rightarrow 0$; that is, C_1 and C_2 are abelian groups generated by three elements and, thus, both are isomorphic to Z^3 (the product of three copies of the group of integers under addition). Now, the homology of the circle will be the quotient groups formed by the kernel of $\P_{n-1}$ mod the image of $\P_n$.

Research indicates that constructing quotient groups is particularly challenging, even among students who have constructed groups as objects (Dubinsky et al., 1994). In the case of computing homologies, there is an additional challenge in making sense of the particular quotient groups defined by a particular homomorphism—the boundary map. Interpreting results geometrically gives these algebraic manipulations a geometric meaning, and the relevant mental actions lie therein. In other words, computing and interpreting homologies becomes a proxy for geometric actions associated with mapping n-dimensional triangles to their boundaries, equating sequences of n-dimensional triangles with an identity element, and forming n-dimensional loops around holes in the topological space under consideration. Thus, we begin to understand the chain complex as a representation of those actions. For the actions to become objectified, we need for them to define a class of spaces, so that homology becomes a proxy for that class.

In taking on this challenge, motivation quickly arises as a competing factor: Why did mathematicians ever bother to invent (co)homology in the first place? This as a competing factor because, for simple examples like the circle, sphere, or torus, there is no need for homology (let alone cohomology). We do not need to compute quotient groups of boundary mappings in order to determine that the torus and the sphere are topologically distinct. On the other hand, for the cases in which homology might be useful, the connection between the topology of the spaces and their homology (roughly, the connection between their geometry and their algebra) is opaque. We need to begin by working with simpler examples in order to build the connection in a way that might extend to ever more complex examples. Along the way, however, new complexities arise within the connection itself.

In working through examples, many of our actions will be conjectural—long sequences of tentative activity with depreciating confidence. For example, we might consider, "Why is homology invariant of choice of simplexes?" After all, we can build up the same topological space in many different ways. As it turns out, we do not even need to use n-dimensional triangles to form a chain complex, but can choose any n-dimensional polygon. Specifically, when computing the homology of the circle, we can choose any number, *m*, as the number of vertices (0-simplices) and edges (1-simplices). Figure 4 illustrates the cases of *m*=1 and *m*=3.

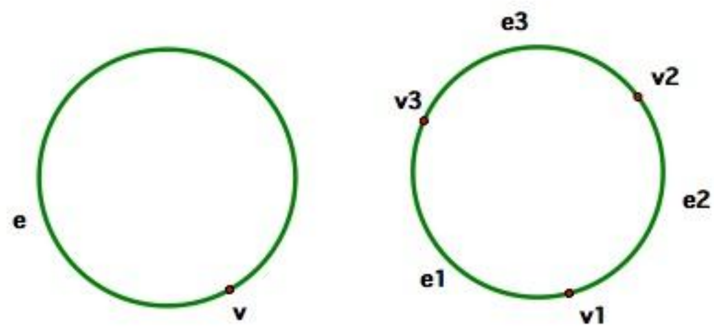


Figure 4. Two ways to form simplexes in the circle.

The image on the right of Figure 4 represents our original approach, with chain complex $0 \rightarrow Z^3 \rightarrow Z^3 \rightarrow 0$. The image on the left generates a simpler chain complex: $0 \rightarrow Z \rightarrow Z \rightarrow 0$. Even though the images and kernels within these mappings differ considerably, the resulting quotient groups are identical. For instance, in computing $H_0(X)$, the corresponding kernels are Z and Z^3, but the corresponding images are 0 and Z^2, so that the quotient group is Z in either case. Understanding why this happens is part of what it means to objectify the quotient groups that define homology. Just as understanding equivalent fractions involves more than showing that common factors cancel out, this understanding relies upon mental actions beyond the computation. Thus, the objectification of homology involves more than an interiorization of the boundary mapping or the process of computing its quotient groups. In particular, every time we add a new vertex to the simplex, we must add another edge, and the boundary of that edge will consist of two adjacent vertices. Their connectivity, as a single connected component, essentially leads to their identification in quotient group: Each vertex is identified with its two adjacent vertices, by the edge that connects them.

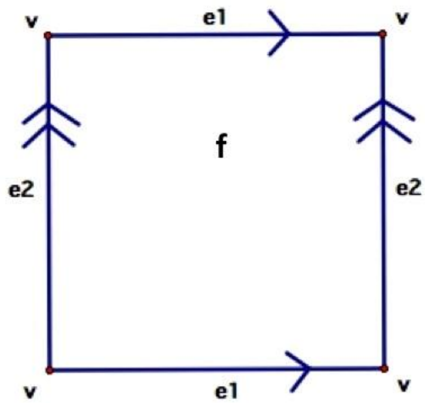
This understanding goes well beyond the process of computing kernels and images of the boundary map, and without this understanding, developed through simple examples, we would not be able to trust the extension of homology to the more complex examples where homology is actually useful. In building an understanding for how the algebraic computation of homology serves as a proxy from making topological distinctions, we find that relevant mental actions include geometric ones, related to vertex-edge graphs, as well as mental actions associated with continuity, especially as it relates to homotopy. By itself, the objectification of the boundary mapping would be no more useful to me than the algorithm for computing the product of two fractions; we would be objectifying something figurative rather than operative, and thus, would not be engaging in mathematics. We dig a little further into these actions by considering two nearly identical surfaces: the torus and the Klein bottle.

The Torus and the Klein Bottle

Topology is intended to address questions like the following: Are the torus and the Klein bottle continuous transformations of one another? Algebraic topology provides an answer by showing that the two surfaces have different homologies. Figure 5 demonstrates the homology of the torus.

$$0 \xrightarrow{\partial_3} < f > \xrightarrow{\partial_2} < e_1, e_2 > \xrightarrow{\partial_1} < v > \xrightarrow{\partial_0} 0$$

$$\partial_1(e) = v - v = 0$$



$$Ker(\partial_0)/\mathrm{Im}(\partial_1) = < v > /0 \cong \mathrm{Z}$$

$$\partial_2(f) = e_1 - e_2 - e_1 + e_2 = 0$$

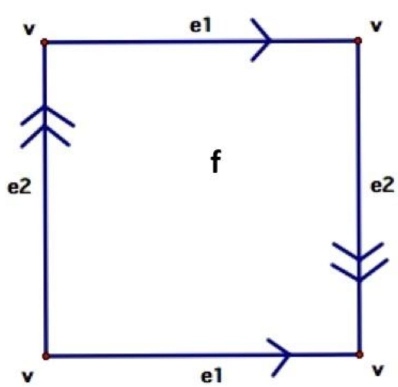
$$Ker(\partial_1)/\mathrm{Im}(\partial_2) = < e_1, e_2 > /0 \cong \mathrm{Z}^2$$

$$Ker(\partial_2)/\mathrm{Im}(\partial_3) = < f > /0 \cong \mathrm{Z}$$

Figure 5. Homology of the Torus

Note that the diagram on the left side of Figure 5 represents a torus because the opposite edges are identified with one another; i.e., we can produce the torus by gluing opposite edges together and, in the process, the four corners become a single vertex, *v*. Also note that each of the boundary maps turn out to be the 0 map because vertices and edges cancel out. Now consider the Klein bottle (Figure 6).

$$0 \xrightarrow{\partial_3} < f > \xrightarrow{\partial_2} < e_1, e_2 > \xrightarrow{\partial_1} < v > \xrightarrow{\partial_0} 0$$



$$Ker(\partial_0)/\mathrm{Im}(\partial_1) = < v > /0 = \mathrm{Z}$$

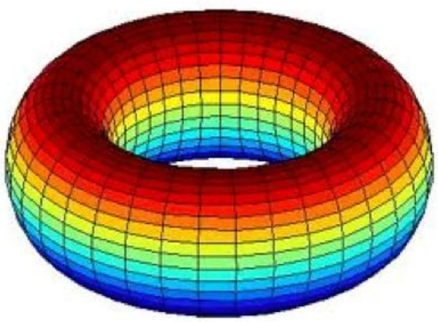
$$\partial_2(f) = e_1 + e_2 - e_1 + e_2 = 2e_2$$

$$Ker(\partial_1)/\mathrm{Im}(\partial_2) = < e_1, e_2 > /2 < e_2 > \cong \mathrm{Z} \times \mathrm{Z}_2$$

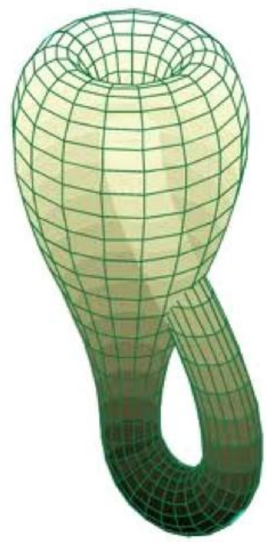
$$Ker(\partial_2)/\mathrm{Im}(\partial_3) = 0$$

Figure 6. Homology of the Klein Bottle.

The diagram (and therefore the homology) is exactly the same, except for one twist: A copy of e_2 is reversed. We can imagine both surfaces being constructed from a cylinder (after the pair of e_1s are identified), but in order to match up the directions of the two copies of e_2, the Kline bottle requires that the cylinder pass through itself to attach from the inside (see right side of Figure 7), which happens in four-dimensional space. Thus, the Klein bottle is a two-dimensional surface that does not exist in three-dimensional space. This fact alone might inform us that the torus and Klein bottle are not topologically equivalent, but we intend the comparison as an explanatory example for homology rather than a motivating one. We are trying to identify mental actions that might underlie our computations.



$$H_0 = Z; H_1 = Z^2; H_2 = Z$$



$$H_0 = Z; H_1 = Z \times Z_2; H_2 = 0$$

Figure 7. Homology as a proxy for topological actions.
Note: Images borrowed from Nosco (http://www.nosco.ch/mathematics/notes/torus.php) and Wikipedia (http://en.wikipedia.org/wiki/Klein_bottle) respectively.

In the case of the circle, we have already seen how the 0^{th} homology group, H_0, indicates the number of connected components in the topological space. Although the torus and Klein bottle affirm this connection (both are connected and have a single copy of Z for H_0), they do not provide interesting cases in this regard because we constructed each of them with only one vertex. However, they do provide an interesting contrast for H_1. How should we interpret the quotient groups Z^2 and ZxZ_2?

For both surfaces, the kernel of the 1^{st} boundary map ($\P_1$) is the group generated by the two edges; both of these edges form loops because their boundary is a single vertex, *v*, and for that same reason, they map to 0. For the torus, those loops are maintained when the face is glued on because the opposite edges match up. In order for them to match up, their directions must be opposite as we go around the boundary, and that is why they cancel out in the 2^{nd} boundary map ($\P_2$). In other words, the 2^{nd} boundary map is 0 precisely because the opposite edges of the face match up. Thus, the image is 0; no paths become identified with 0 in the quotient; and the 1^{st}

homology group (H_1) is the group generated by the two loops. We can see these loops on the torus in Figure 7: One goes around the "inner tube" and one goes around the hole at the center of the torus.

For the Klein bottle, one of the loops is transformed when the face is glued on because one pair of opposite edges does not match up. Instead of canceling out, the edge is doubled, and the 2nd boundary map has an image of $<2e_2>$. Thus, any even number of trips around the corresponding loop will be identified with 0. We can see the corresponding geometry in Figure 7: Tracing a loop around the "neck" of the bottle is just as it was for the "inner tube" of the torus, but tracing the other way yields a loop that undoes itself on the second pass because the trace moves to the other side of the surface (from inside out, or vice versa).

In general, the *kernel* of a boundary map is generated by n-dimensional *cycles,* and the *image* of the next boundary map is generated by the n-dimensional *boundaries* of n+1 dimensional polygons. In fact, algebraic topologists refer the kernels and images as "cycles" and "boundaries," respectively. In the quotient groups that define homology, the boundaries are identified with 0. Geometrically, we can understand this as gluing the cycles together (often in intricate ways). However, we can get lost in the computation of cycles, boundaries, and their quotients without ever considering the geometric actions to which they refer, much as middle school students do when they "complete the square" without ever considering the geometric square they are completing. Whether we are completing squares, connecting vertices, or gluing faces on to loops, the mathematics is in the geometric action for which the algebraic manipulation is a proxy. Once these actions are objectified, they can be symbolized in a way that conveys meaning. In particular, the symbols in Figure 3 become more than figurative material; they become proxies for objects, and actions on those objects.

Concluding Remarks

In reflecting on the actions and objects of cohomology, a key distinction arises—one that Piaget vigilantly maintained in his studies of young children but one that becomes easier to overlook when considering advanced mathematics: The bases for construction of formal mathematical objects are not necessarily formal processes. The diagram presented in Figure 3 might implicate computing kernels and images of boundary maps as primary actions to objectify, but subsequent investigation indicates a wide network of mostly geometric actions to coordinate. This finding supports the Piagetian notion that mathematics is a product of *psychological* action and not simply the enculturation of formal processes developed in the history of mathematics.

APOS theory (Dubinski, 1991) and reification (Sfard, 1992) have contributed greatly to mathematics education by extending Piaget's notion of reflective abstraction to advanced mathematics. However, researchers tend to use these frameworks as pedagogical tools for supporting student mastery of formal procedures, such as computing quotient groups (Asiala, Dubinsky, Matthews, Morics, & Oktac, 1997), especially when *actions* and *processes* refer to formal procedures. Although computations and procedures are integral to mathematical development, we must explicitly attend to the mental actions that give them meaning in order to support operative (and therefore mathematical) knowledge, rather than figurative knowledge. In fact, Sfard herself pointed to the "pitfall" of figurative knowledge when she warned of *pseudostructuralist* approaches to knowledge and learning (Sfard & Linchevski, 1994), which are indicated in students' aversion to "procedures with any degree of freedom" (Leron, Hazzan, & Zazkis, 1995). In contrast, a structuralist approach to mathematical knowledge and learning focuses on the construction and organization of reversible mental actions (Piaget, 1970b).

Scheme theory (von Glasersfeld, 1995) adopts a structuralist approach but has its own limitations in modeling the development of advanced mathematics; namely, the simplicity of a three-part structure may not accommodate the complexity of advanced mathematical concepts. Although we are able to identify some of the mental actions that undergird cohomology, we do not have models for their organization. This may explain why we often revert to figurative representations of knowledge (e.g., Figure 3) when investigating the development of advanced mathematics.

Our investigation of cohomology supports the argument that mathematics, at all levels, can be characterized as the objectification of action. This is the defining feature of mathematics, which distinguishes it from all other languages and sciences. Understanding mathematics in this way also evokes a degree of empathy as we provoke our students to construct new objects through action. In Bill Thurston's case, the father did not appreciate his son's accomplishment in constructing improper fractions as "numbers in their own right" (Hackenberg, 2007) because he could not unpack the coordinated actions of that construct. Likewise, models for teaching and learning advanced mathematics are limited by our models of the mental actions that comprise the objects of advanced mathematics.

References

Asiala, M., Dubinsky, E., Matthews, D. M., Morics, S. & Okatc, A. (1997). Development of students' understanding of cosets, normality, and quotient groups. *The Journal of Mathematical Behavior*, *16*(3), 241-309.

Brown, A., DeVries, D. J., Dubinsky, E., & Thomas, K. (1997). Learning binary operations, groups, and subgroups. *The Journal of Mathematical Behavior*, *16*(3), 187-239.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95-123). Netherlands: Kluwer Academic Publishers.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics*, *27*(3), 267-305.

Dubinsky, E., & Lewin, P. (1986). Reflective abstraction and mathematics education: The genetic decomposition of induction and compactness. *Journal of Mathematical Behavior, 5*, 55-92.

Dubinsky, E., & McDonald, M. A. (2001). APOS: Constructivist theory of learning in undergraduate mathematics education research. *The Teaching and Learning of Mathematics at University Level, 7*(3), 275-282.

Hackenberg, A. J. (2007). Units coordination and the construction of improper fractions: A revision of the splitting hypothesis. *The Journal of Mathematical Behavior, 26*(1), 27–47.

Hazzan, O. (1999). Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics, 40*, 71-90.

LaBoskey, V. K. (2004). The methodology of self-study and its theoretical underpinnings. In J. J. Loughran, M. L. Hamilton, V. K. LaBoskey, & T. Russell (Eds.), *International handbook of self-study of teaching and teacher education practices* (pp. 817–869). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Leron, U., Hazzan, O., & Zazkis, R. (1995*)*. Learning group isomorphism: A crossroads of many concepts. *Educational Studies in Mathematics, 29, 153-174.*

Norton, A., & Wilkins, J. L. M. (2012). The splitting group. *Journal for Research in Mathematics Education, 43*(5), 557-583.

Piaget, J. (1970). *Genetic epistemology* (E. Duckworth, Trans.). New York: Norton.

Piaget, J. (1970). *Structuralism* (C. Maschler, Trans.). New York: Basic Books.

Piaget, J. (1971). *Psychology and epistemology* (A. Rosin, Trans.). New York: Grossman Publishers (Originally published in 1970).

Piaget, J., & Garcia, R. (1991). *Toward a logic of meanings.* Hillsdale, NJ: Erlbaum.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on process and objects on different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1-36.

Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification – the case of function. In E. Dubinsky & G. Harel (Eds.), *The Concept of Function: Aspects of Epistemology and Pedagogy* (pp. 59-84). Washington, Mathematical Association of America.

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification—the case of algebra. In *Learning mathematics* (pp. 87-124). Springer Netherlands.

Steffe, L. P., & Olive, J. (Eds.) (2010). *Children's fractional knowledge*. New York: Springer.

Tall, D., Thomas, M, Davis, G., Gray, E., & Simpson, A. (2000). What is the object of an encapsulation of a process? *Journal of Mathematical Behavior, 18*(2), 223-241.

Thompson, P. W. (1985). Experience, problem solving, and learning mathematics: Considerations in developing mathematics curricula. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 189–243). Hillsdale, NJ: Erlbaum.

Thurston, W. P. (1990). Mathematical education, *Notices of the American Mathematical Society, 37*, 844-850.

Valéry, P. (1974). *Cahiers, Vol. II*. Paris: Gallimard, 1974.

von Glasersfeld, E. (1995*). Radical constructivism: A way of knowing and learning*. London: Routledge Falmer.

15

INSTRUCTIONAL PRACTICES AND STUDENT PERFORMANCE IN CALCULUS

Lisa Mantini	Barbara Trigalet	R. Evan Davis
Oklahoma State University	Texas Academy of Biomedical Sciences	Oklahoma State University

Classroom teaching in multiple sections of Calculus I at a large comprehensive research university was observed and coded using the Teaching Dimensions Observation Protocol (TDOP). Multiple teaching styles were identified ranging from low engagement to moderate engagement to high engagement sometimes including student group work. Student performance on two course-wide uniform exams and on the Calculus Concept Inventory (CCI) was analyzed for any correlations with teaching methods. Significant correlations were found between high engagement teaching styles and performance on both the first exam and the final exam. However, section normalized gains on the CCI were found to be significantly correlated only with the presence of student group work or desk work and with no other measures of teaching practice or student performance.

Key words: [Calculus instruction, classroom observations, student performance, calculus concepts inventory, teaching dimensions observation protocol]

Introduction and Literature Review

The United States is not producing enough graduates in Science, Technology, Engineering and Mathematics (STEM) (Bressoud, 2011) and the need is particularly great in the mathematically intensive majors. However, college freshmen entering one of the STEM majors face a significant hurdle in Calculus I. Currently, the Mathematical Association of America is investigating the teaching of college calculus courses nationwide to describe and measure the impact of the various characteristics of calculus classes that appear to influence student success (Bressoud et al., 2013; Rasmussen et al, 2014). As Speer, Smith and Horvath note, "research on collegiate teachers' actual classroom teaching practice is virtually non-existent" (2010, p. 99). According to Bressoud (2012), "the mathematical community does not have research evidence for instructional strategies that work." This study seeks to contribute to a growing body of research on actual classroom practice, as well as determine possible correlations between actual classroom practices and student achievement in university courses in calculus.

While much research has found alternatives to lecture such as "inquiry-oriented" or "constructive process" pedagogies to be successful (Ganter, 1999; Rasmussen, Kwon, Allen, Marrongelle & Burtch, 2006; Kogan & Laursen, 2013), others have found lecture to be effective (Hora & Ferrare, 2013; Saroyan & Snell, 1997) or preferable to students (Ferrini-Mundy & Güçler, 2009; Murray, 1983). This suggests that there is a need for a detailed description of in-class instruction to capture the relations among instructors, students and classroom environments. Porter (2002) notes that careful analysis of teaching can help identify methods that contribute to student achievement.

A growing trend in the assessment of student understanding is the use of Concept Inventories, dating back to the work in physics of Halloun and Hestenes (1985) in developing the Force Concept Inventory (FCI). The FCI is intended to serve as a reproducible and objective measure of how a course improves comprehension of principles (Epstein & Yang, 2007); higher gains are seen after interactive engagement pedagogies in which students receive immediate feedback in class on their understanding of a topic. Similarly, the Calculus Concept Inventory (CCI) (Epstein, 2013) purports to measure conceptual understanding of the principles of calculus through the use of multiple choice questions requiring little to no calculation. Typically the CCI is given as a pre-test and post-test in one semester, and sections of the course are compared by comparing their normalized gain, which is the ratio of actual gain in the class average score (post-test mean less pre-test mean) divided by maximum possible gain (maximum possible score less pre-test mean), though other measures of score comparison are possible (Thomas & Lozano, 2013).

Research Questions

The research questions addressed by this study are:

1. What instructional practices including teaching methods, pedagogical moves, instructor/student interactions, cognitive engagement and instructional technology are being used in Calculus I at a large research university?
2. (a) Which of these practices correlate to increased student conceptual understanding as measured by normalized gain on the Calculus Concepts Inventory? (b) Which of these practices correlate to higher average student performance on a uniform final exam?

Methods

Setting

At the large, comprehensive research university during the term when this study took place, Calculus I was taught in small sections with from 36 to 43 students per section. These sections met for either four 50-minute meetings or three 75-minute meetings per week; class start times ranged from 8:00 AM until 2:30 PM. Section enrollments were unrestricted, and students self-enrolled into their preferred section. Course coordination was handled by a member of the tenured faculty. The common elements included the syllabus, online homework, first exam, and final exam, with both common exams graded uniformly. Other items such as the second and third hourly exams, any written homework, quizzes, gateway testing, and group work were determined individually by the instructors.

Traditionally at this university, spring enrollments in Calculus I are smaller than in the fall and DWF rates are higher than in fall semesters. A significant proportion of students taking Calculus I in the spring had been placed into a pre-calculus course in the fall and then passed that course with a grade of C or better. Another significant proportion of the spring Calculus I population consisted of students who were retaking Calculus I after an unsuccessful experience in Calculus I in the fall.

Participants

The instructors who participated in the study consisted of 10 volunteers from among the section instructors, responsible for 11 sections of Calculus I (one volunteer was teaching 2 sections). Two instructors were tenured professors with substantial teaching experience; four of the remaining instructors had held the Ph.D. for four years or less, and the remaining four instructors were advanced doctoral students within a year or two of earning the Ph.D. Four instructors were teaching their own section of Calculus I for the first or second time; all others had prior experience as an independent instructor in Calculus I. Four instructors were in their first year of teaching at the study institution. Six of the instructors were American and four were internationals; eight were male and two were female. Study participants accounted for over 90% of the sections of Calculus I taught during the semester in question and over 90% of the students enrolled in Calculus I during that semester.

Student participants were solicited in the sections taught by all instructors who were participating in the study. Phase I participants consisted of the 347 volunteers who completed both uniform exams, Exam 1 and the Final Exam. These students accounted for over 70% of students enrolled in Calculus I that semester. Phase II participants were the 208 volunteers who also completed both the pre-test and the post-test for the Calculus Concepts Inventory.

These students essentially consisted of 60% of the Phase I participants; they represented from 34% to 72% of students from each section participating.

Data Collection

The uniform course exams were written by the coordinator with input from all of the section instructors. Questions were fairly standard and emphasized calculations but included some conceptual questions, some real-world applications, and some items requiring multiple representations of functions such as determining information about a function from the graph of its derivative. All items were free response questions except for one short answer question. Grading was done uniformly, with one instructor grading one problem on all papers. Scores for student study participants were reported to the researchers.

The Calculus Concepts Inventory was administered in class once during week 1 and once during week 15 by all instructors. This is a multiple choice instrument requiring little calculation which tests the student's understanding of calculus concepts (Epstein, 2012). In some cases the instructors scored their own sections and reported results to the researchers; otherwise the papers were scored by the researchers. Attendance was highly variable on the days when the CCI was administered.

Classroom observations were conducted using the Teaching Dimensions Observation Protocol (TDOP) (Hora and Ferrare, 2010). This instrument codes which of multiple behaviors by teachers or students are observed during each 2-minute interval of an observation. It has been used previously to classify instructional behaviors in college-level instruction in Calculus (Code, Kohler, Piccolo, and MacLean, 2012) and across disciplines (Hora and Ferrare, 2013). Instructors participating in the study had access to the instrument and were aware that the broad categories being observed were Teaching Methods, Pedagogical Moves, Instructor-Student Interaction, Cognitive Engagement, and Instructional Technology (see Appendix for a table listing all TDOP codes). Each section in the study was observed 3 times, during month 1, month 2, and month 4. Before each observation, the observer contacted each instructor to ascertain that the observed class period would be what the instructor would call "typical." All observations were done live, not from video, and by one researcher only. Before using the instrument for live observations, however, several researchers practiced coding from video and compared results, in order to train themselves on using the instrument and to increase inter-rater reliability.

Other observation instruments were considered and rejected for this study. Among these were the Teacher Behavior Inventory (TBI) (Murray, 1983), which gathers subjective accounts from students assessing instructor behaviors, and the Reformed Teaching Observation Protocol (RTOP) (Sawada et al, 2002), which aims to evaluate the extent to which instruction meets the goals of being inquiry-oriented or student-centered, and thus does not provide a descriptive account of teaching behaviors (Hora and Ferrare, 2013).

Data Analysis

For each of the 11 sections in the study, observational data from the TDOP were converted into a sequence of 0's and 1's, where a 1 was recorded if that particular behavior was observed in a two-minute interval and a 0 if not. These data were entered into an Excel spreadsheet. Each section was observed 3 times, so the total number of observed 2-minute intervals ranged from 71 to 114 per section (some sections met for 50 minutes, some for 75 minutes, and class periods occasionally ended early or ran a bit long). We then determined the proportion of observed 2-minute intervals in which each particular TDOP code was observed. This gave us a range of proportions for each TDOP code indicating its relative frequency of use among study participants. Many codes varied little across sections, but those codes that had high variability across sections were noted.

For Phase I of the analysis, student performance was averaged in each section, producing two data points summarizing student performance: the final exam average and the CCI normalized gain. CCI normalized gain is computed as the ratio of the actual section mean gain (post-test mean less the pre-test mean) to the maximum possible mean gain (maximum score minus the pre-test mean), that is,

CCI normalized gain = (post-test mean – pre-test mean)/(22 – pre-test mean).

Pearson correlation coefficients were computed between each of the student performance indicators and TDOP proportions across sections. TDOP categories showing a significant correlation with student performance were noted.

In Phase II of the analysis, aggregate codes were formed from related TDOP categories with significant correlation to student performance measures, in order to try to broadly categorize instructional practices and student engagement. As a result, we were able to form an instructional profile for each section in the study and to characterize these into three groups: Group 1 consisted of sections displaying low engagement, Group 2 consisted of sections displaying moderate engagement, and Group 3 consisting of high engagement sections. Using these instructional groups, we then created a spreadsheet of anonymized individual student scores on four measures: the CCI pre-test score, the exam 1 grade, the CCI post-test, and the final exam score, along with the group number indicating the engagement level of that student's instructor. Additional analysis was performed including ANOVA and ANCOVA to determine if any correlation was present between the instructional profile and student performance.

Results

Teaching Practices

Initial findings from the TDOP regarding Teaching Methods indicate that all instructors employ lecturing with visuals, seen in 80% of the two-minute intervals coded. The instructional technique of having students work at their desks, either in small groups (SGW) or by themselves (DW), was observed 11% of the time but it was used by only four instructors, ranging from 16% to 32% of the time in those sections. Several codes in both the Instructor-Student Interaction category and the Cognitive Engagement category varied significantly. Overall, approximately 60% of time intervals coded contained questions asked of the students by the instructors, with students responding in more than 50% of the time intervals coded. However, some instructors used display questions (DQ), asking students to display content knowledge, as often as 85% of the time, others as little as 11% of the time. Among Instructional Technology, the most predominant tool was the chalk board or white board, used 77% of the time. The use of power point slides and a digital tablet varied significantly, ranging from no use to use more than 30% of the time. Other instructional technologies were observed well less than 10% of the time.

Student Performance

Final exam scores were available for 347 student participants. The average final exam score for the participants was 68.5%. The final exam averages for sections participating in the study ranged from a low of 55% to a high of 75%, with two sections having averages between 55.2% and 62.2%, two sections with averages in between 66.2% and 67%, and seven sections with averages in between 69.8% and 75.2%.

CCI scores were available for only 208 of 347 students completing the course, or 60% of the students who took the final exam. The CCI normalized gain among all students in the study was 10.2%. This is a fairly low number for normalized gain as compared to reports of other studies (Thomas & Lozano, 2013) and as compared to our data from a prior semester in which we obtained a course wide normalized gain of 17% (data analysis is ongoing). The individual sections in the study had normalized gains on the CCI ranging from 3.4% to 20.3%, with three sections having gains in between 3.4% and 3.7%, two sections with gains from 9.6% to 9.7%, four sections with gains from 12.1% to 13.3%, and two sections with a gain between 15.1% and 20.3%. However, it is notable that not all sections had the same rate of participation in the CCI. Attendance was highly variable on the day when the CCI post-test was given, and in many cases the student population from which the CCI data is drawn may include a disproportionate number of stronger students. The proportion of section enrollment completing both the CCI pre-test and post-test ranged from a low of 34.4% to a high of 72% in this study, with two sections having fewer than 38% of students participating, seven sections having between 45% and 59% of students participating, and two sections having between 62% and 72% of students participating. It is notable that the highest CCI normalized gains occurred in sections with fewer than 55% of students participating.

Aggregate Codes and Instructional Profiles

For the eleven sections in the study, we computed Pearson's correlation coefficients for all of the TDOP codes with both the section average final exam score and with the CCI normalized gain. Several codes showed significant correlations. The code SGW, student group work, was positively correlated with CCI normalized gain. The code DW, individual desk work, was positively correlated with the final exam average. Code ART, student articulation, was positively correlated with the final exam average as well. Code A, assessments, had a negative correlation with the final exam average, and the code DT, indicating use of a digital tablet or document camera with visuals prepared before class, had a negative correlation with the final exam average performance.

Based on this preliminary analysis, we combined several related TDOP codes to create aggregate codes in order to search for stronger positive correlations with the student performance measures. Teaching methods were observed in 93% of the observed time intervals and at least 85% of the time in all sections. The aggregate code SWK indicates the proportion of time when a teaching method was observed and students were observed actively working on problems via code SGW or DW. The aggregate code SVB indicates the proportion of time when a teaching method was observed and students were observed answering questions posed by the instructor, asking questions, or other forms of articulation (see list of all TDOP codes in Appendix). The sum of these two codes is abbreviated as SENG, to indicate that students were seen to be actively engaged in either of these manners. Finally, the code LNWV indicates the proportion of all time when a teaching method was observed but students were not seen to be either working or verbalizing. These codes are described in Table 1.

Table 1: Aggregate TDOP Codes

Code	Meaning	Description
TMTH	Teaching method observed	Coded as 1 when any teaching method is observed
SWK	Students working	Coded as 1 when either SGW (student group work) or DW (individual desk work) is coded
SVB	Students verbalizing	Coded as 1 when SWK = 0 and a student response or question is coded (SNQ, SCQ, SR, or ART)
SENG	Students engaged	Coded as 1 when TMTH = 1 and either SWK = 1 or SVB = 1
LNWV	Lecture, no work or verbalization	Coded as 1 when TMTH = 1 and SENG = 0

The resulting instructional profiles indicate a range of instructor behaviors. Code SWK ranged from 0% to 31.9%; code SVB ranged from 18.8% to 81.6%; and SENG, the sum of

SWK and SVB, ranged from 21.9% to 88.2%. As a result, LNWV ranged from a high of 78.1% down to a low of 11.8%. Note that SENG + LNWV = 100%. The data suggested sorting instructor profiles into three groups. It is notable that one instructor in this section taught two sections, and the two sections were assigned to different instructional profiles.

Table 2: Instructional Profiles Observed

Group	Engagement Profile	SENG values	No. sections
1	Low	20% - 40%	4
2	Moderate	50% - 60%	3
3	High	80% - 90%	4

Figure 1: Illustration of Observed Instructional Profiles

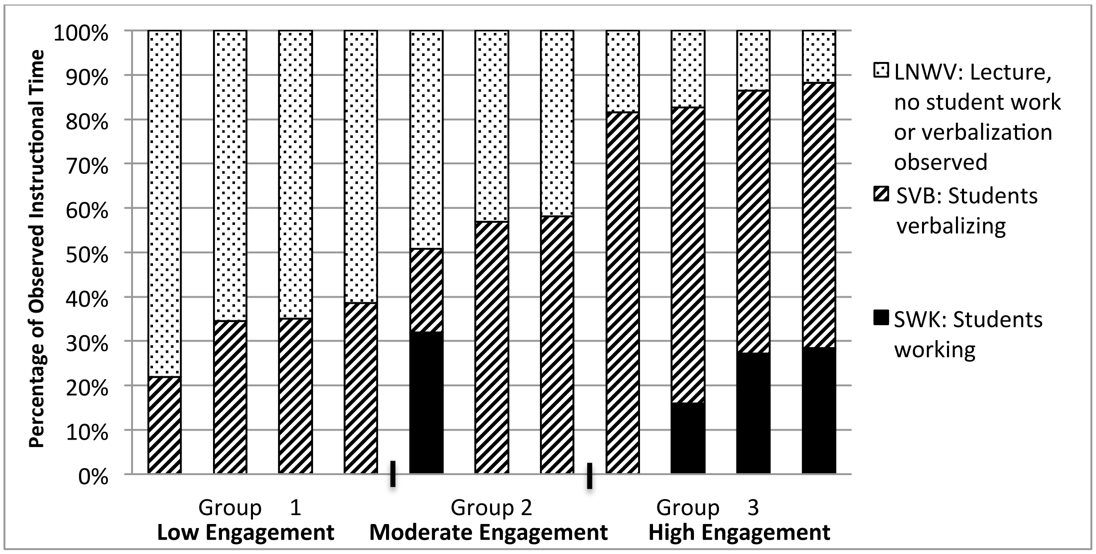


Figure 1 illustrates the range of instructional profiles seen in these observations. The teaching profile bars are arranged from left to right by increasing SENG proportion, and the three instructional groups are indicated.

Correlations of Instructional Profiles with Section Average Performance

Using our data from the 11 course sections in the study, we calculated Pearson's correlation coefficients between our aggregate TDOP codes SWK, SVB, and SENG and average student scores in each section on four measures: Exam 1, the Final Exam, the Sum of Exam 1 and the Final Exam, and the section's net gain on the CCI. Results indicated that the Exam 1 average was correlated significantly ($p < .05$) with code SENG. The Final Exam average and the Exam Average Sum were both correlated significantly ($p < .05$) with code SWK and correlated highly significantly ($p < .01$) with the combined code SENG. The section CCI normalized gain was correlated significantly ($p < .02$) with the code SWK but not correlated with either SENG or SVB.

Figure 2 illustrates the instructional profile bars in order by final exam average in each section, with the placement of the bars corresponding to final exam average. Notice that the graph illustrates the correlation of the final exam average with the SENG code (represented by the combination of solid and striped areas).

Figure 2: Instructional Profiles Ordered by Final Exam Average

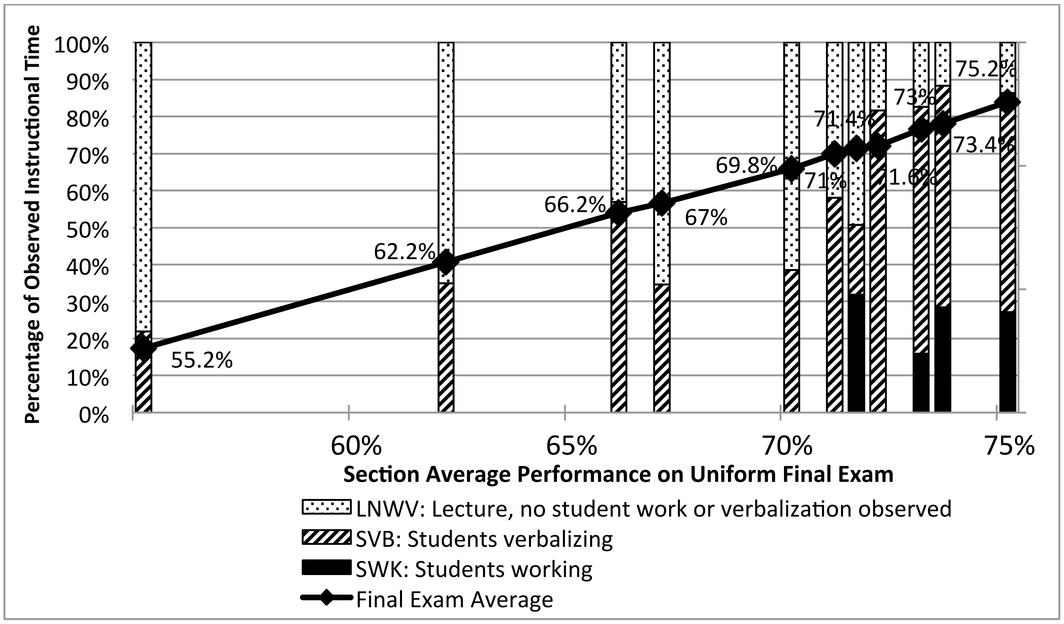
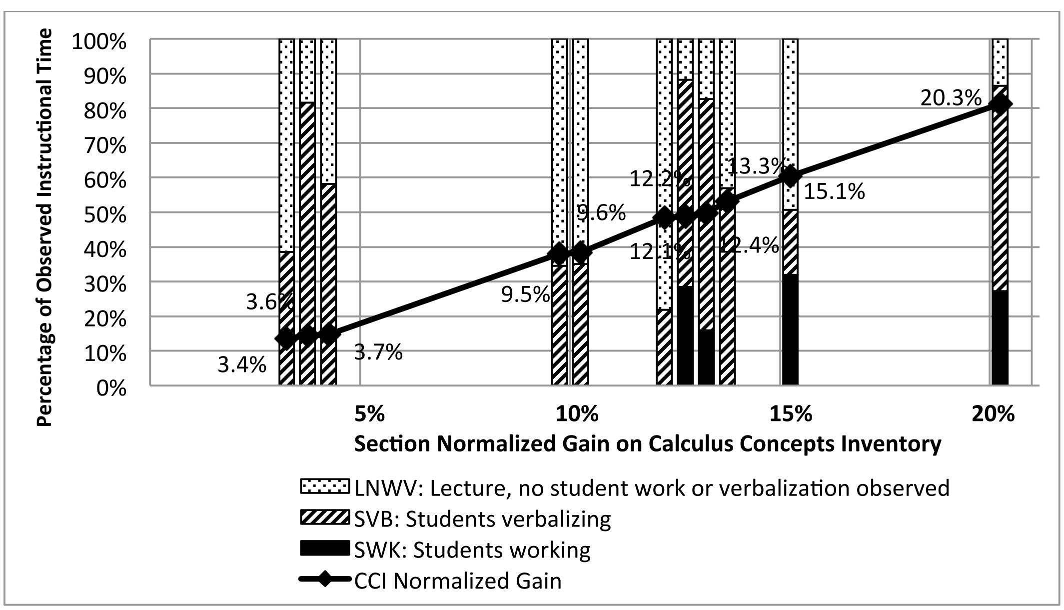


Figure 3 below illustrates the instructional profile bars in order by CCI normalized gain in each section, with the placement of the bars corresponding to the CCI normalized gain. Notice that the graph illustrates that the CCI normalized gain is correlated significantly with code SWK, students working, but is uncorrelated with codes LNWV and SVB.

Figure 3: Instructional Profiles Ordered by CCI Normalized Gain



Statistical Results for Individual Student Performance

The student data consisted of scores on each of four assessment measures along with a variable which sorted student scores into three groups according to the instructional profile assigned to their instructor, with 1 indicating the low engagement profile, 2 indicating moderate engagement, and 3 indicating high engagement. The assessment measures considered were the CCI Pre-test, administered in week 1; the score on Exam 1, administered in week 5, the CCI Post-test, administered during week 15; and the uniform Final Exam, administered during week 16 of the semester. We used SPSS software to search for any significant correlations of student performance on the various assessments with the instructional profiles assigned.

We found no significant difference [F(2,305)=1.88, n.s.] among students across the three groups in the analysis of the CCI pre-test.

When comparing the scores on common Exam 1, analysis of variance revealed a significant difference in performance [F(2,347) = 12.84, p < .01] among students. Examination of paired comparisons (Tukey and Scheffe) showed that, while the moderate and low engagement groups did not significantly differ from each other, the high engagement group scored significantly better on the first exam than either of the other two conditions.

When comparing the scores on the common Final Exam, analysis of variance again revealed a significant difference across the three groups [F(2, 347) = 7.46, p < .01]. Paired comparison between the three groups revealed that, while the difference between the high and low engagement groups was still significant, the difference between the moderate engagement group and either the low engagement group or the high engagement group was not statistically significant. The results for the final exam are interesting in that they suggest that the moderate engagement group "gained ground" on the high engagement group between the first common exam and the final exam, with a higher estimated marginal mean for the moderate engagement group as compared to the high engagement and low engagement groups in an ANCOVA analysis with the final exam as our dependent variable and the first exam as covariate.

We also found no significant difference [F(2,215)=.08, n.s.] among students across the three groups in the analysis of the CCI Post-test.

A summary of means on the student performance measures in each group is included in Table 3.

Table 3: Means of Student Performance Measures Across Instructional Groups

Measure	Group 1 Mean	Group 2 Mean	Group 3 Mean
CCI Pre-test raw score	6.81	6.49	6.49
percentage	31%	29.5%	29.5%
CCI Post-test raw score	8.16	8.21	8.29
percentage	37.1%	37.3%	37.7%
CCI normalized gain	8.9%	11.1%	11.6%
Uniform Exam 1	70.8%	69.3%	79.5%
Uniform Final Exam	63.2%	69.6%	73.2%

Discussion

Regarding our first research question, we found that the teaching methods observed relied primarily on lecture methods, seen from 68% to 100% of the time. Within lecture methods, though, the use of questioning and other engagement techniques varied significantly. Our data seem to indicate a possible definition of high engagement instruction, but further research is needed. It is interesting to note that all instructional profile groupings included instructors of varying experience levels and both Americans and internationals.

Regarding our second research question, the correlation of section normalized gain on the CCI with code SWK, agrees with some prior results reported in the literature (Epstein, 2013;

Thomas & Lozano, 2013) but this is difficult to understand fully. The number of students participating in both CCI pre-test and post-test (n=208) is small, as little as 34% of enrollment in some sections, and may contain the better students in each section, since many of those absent on the days when the CCI was administered may have been weaker students. The lack of correlation between CCI pre-test and post-test scores and any TDOP variables or other assessments bears further investigation but may indicate a lack of effort by participants on the CCI, which did not count towards their course grade. Using additional computational methods to analyze CCI scores such as individualized gain scores or item response theory (Thomas & Lozano, 2013) may shed additional light on the relation of the CCI scores to observed instructional practices and to other student performance measures.

The high correlation of exam scores with the level of engagement in the instructional profile is very interesting and also deserves further study. This result indicates that there may be a benefit derived from providing training to new Calculus I instructors in effective use of questioning techniques and of group work, and of supporting more experienced instructors who wish to adopt these effective instructional strategies. Of course, many other variables affect student performance in Calculus I, including the student's background and preparation in algebra and pre-calculus. Other studies (Bagley, 2014) have shown that it is possible for variations in student understanding of pre-calculus concepts to account for all of the variation seen in student performance across sections in Calculus I. Additional analysis of our data would be desirable to determine how much of the performance variation we are seeing may be accounted for by differences in background and preparation.

Further research is desirable to investigate if there is any correlation between teaching methods and persistence in the calculus sequence or student performance in later courses. Studies of inquiry-based instructional practices (Kogan & Laursen, 2013) have shown increased persistence in mathematics courses by students experiencing inquiry-based instruction as opposed to traditional lecture instruction, and these effects were shown to be sizable and persistent with previously low-achieving students. Longitudinal data may shed light on whether the differences in lecture-based teaching methods that we have observed influence student persistence in the calculus sequence or persistence in STEM majors. More observational data might provide richer descriptions of teaching styles in use in Calculus I and further evidence to support the correlations we found. Interviews with instructors might shed light on their decisions with regard to engagement levels and could be relevant to instructional training programs.

References

Bagley, S. (2014). A comparison of four pedagogical strategies in calculus. In T. Fukawa-Connelly, G. Karakok, K. Keene, and M. Zandieh, (Eds.), *Proceedings of the 17th Annual Conference on Research in Undergraduate Mathematics Education*, Denver, Colorado. In press.

Bressoud, D. (2011, February 1) Status of the math-intensive majors. [Web log post]. Retrieved from http://www.macalester.edu/~bressoud/pub/launchings /launchings_02_11/launchings_02_11.html#anchor1

Bressoud, D. (2012, July 1) Learning from the physicists. [Web log post]. Retrieved from http://launchings.blogspot.com/2012/07/learning-from-physicists.html

Bressoud, D., Carlson, M., Mesa, V., & Rasmussen, C. (2013). The calculus student: Insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology*, *44*, 1-15.

Code, W., Kohler, D., Piccolo, C., & MacLean, M. (2012). Teaching methods comparison in a large introductory calculus class. In S. Brown, S. Larsen, K. Marrongelle, & M. Oehrtman, (Eds.) *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education,* (pp. 375-379), Portland, Oregon.

Epstein, J. & Yang, D. (2007). Development and validation of the calculus concept inventory. In D. Pugalee, A. Rogerson, & A. Schinck, (Eds.), *Proceedings of the Ninth International Conference on Mathematics Education in a Global Community,* (pp. 165-170). Charlotte, NC.

Epstein, J. (2013). The calculus concept inventory – Measurement of the effect of teaching methodology in mathematics. *Notices of the American Mathematical Society*, *60*, 1018-1025.

Ferrini-Mundy, J., & Güçler, B. (2009). Discipline-based efforts to enhance undergraduate STEM education. *New Directions for Teaching and Learning*, *117*, 55-67.

Ganter, S. L. (1999). An evaluation of calculus reform: A preliminary report of a national study. In B. Gold, S. Keith, & W. Marion (Eds.), *Assessment Practices in Undergraduate Mathematics*, (pp. 233-236). Washington, DC: Mathematical Association of America Notes.

Halloun, I. & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics, 53*(11), 1043-1048.

Hora, M., & Ferrare, J.. (2010). The Teaching Dimensions Observation Protocol (TDOP). Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Hora, M. T. & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of undergraduate math and science course planning and classroom teaching. *Journal of the Learning Sciences*, *22*(2), 212-257.

Kogan, M. & Laursen, S. (2013). Assessing Long-Term Effects of Inquiry-Based Learning: A Case Study from College Mathematics. *Innovative Higher Education*, retrieved from http://link.springer.com/article/10.1007/s10755-013-9269-9/fulltext.html.

Murray, H.G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology*, *75*, 138-149.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, *31*(7), 3-14.

Rasmussen, C., Ellis, J., & Zazkis, D. (2014). Lessons learned from case studies of successful calculus programs at five doctoral degree granting institutions. In T. Fukawa-Connelly, G. Karakok, K. Keene, & M. Zandieh, (Eds.), *Proceedings of the 17th*

Annual Conference on Research in Undergraduate Mathematics Education, Denver, Colorado. In press.

Rasmussen, C., Kwon, O. N., Allen, K., Marrongelle, K. & Burtch, M. (2006). Capitalizing on advances in mathematics and K-12 mathematics education in undergraduate mathematics: An inquiry-oriented approach to differential equations. *Asia Pacific Education Review*, *7*(1), 85-93.

Saroyan, A. & Snell, L.S. (1997). Variations in lecturing styles. *Higher Education*, *33*, 85-104.

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R. and Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics*, *102*, 245-253.

Speer, N. M., Smith III, J. P. & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99-114.

Thomas, M. & Lozano, G. (2013). Analyzing Calculus Concept Inventory gains in introductory calculus. In S. Brown, G. Karakok, K. H. Roh, & M. Oehrtman, (Eds.), *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education*, (pp. 637-646), Denver, Colorado.

Appendix: TDOP Codes

This table contains all of the codes used in the TDOP.

Table 4: Codes Used in the Teaching Dimensions Observation Protocol

	Teaching Methods		**Pedagogical Moves**
L	Lecture, no visuals	MOV	Moves into audience
LPV	Lecture, pre-made visuals	HUM	Humor
LHV	Lecture, handwritten visuals	RDS	Reads verbatim from notes or text
LDEM	Lecture with demonstration	IL	Illustration from real world
LINT	Interactive lecture	ORG	Organization
SGW	Small group work	EMP	Emphasis
DW	Desk work	A	Assessment
CD	Class discussion	AT	Administrative task
MM	Multimedia		**Instructional Technology**
SP	Student presentation	PO	Posters used
	Instructor/Student Interaction	B	Books used
RQ	Instructor rhetorical question	N	Lecture notes actively used
DQ	Instructor display question	P	Pointer used
CQ	Instructor comprehension quest.	CB	Chalk board or white board used
SNQ	Student novel question	OP	Overhead or transparencies used
SCQ	Student comprehension quest.	PP	Powerpoint or digital slides used
SR	Student response	CL	Clickers used
	Cognitive Engagement	D	Demonstration equipment used
ART	Articulation by students	DT	Digital tablet or document camera used
RMF	Reciting or memorizing facts	M	Movie, documentary, other video clip

PS	Problem solving	SI	Simulation
CR	Students create their own ideas	WEB	Website or other online resource
CN	Connections to real world	OB	Object used as part of instruction

A FRAMEWORK AND A STUDY TO CHARACTERIZE A TEACHER'S GOALS FOR STUDENT LEARNING

Frank S. Marfai
Arizona State University

In this study, a secondary school teacher's goals for student learning were characterized using a framework that emerged from prior work. Observed lessons spanning the use of both conceptually rich and skill-based curricula were analyzed. The findings suggest that both challenges and opportunities exist for professional development endeavors that center around perturbing a teacher's goals.

Key words: Teacher Goals, Mathematical Knowledge for Teaching, Teacher Knowledge, Teacher Beliefs, Professional Development

It is widely known that mathematics teaching in the United States has been characterized as procedural and disconnected (Ma, 1999; Stigler & Hiebert, 1999), with little focus on understanding how mathematical concepts develop and how they are connected. In recent work it has also been documented that it is common for teachers to teach in a manner in which they were instructed as students, and that making the transition to value conceptual learning and teaching is a difficult transition for teachers to make (Sowder, 2007).

Theoretical Framework

Researchers have identified mathematical knowledge for teaching as a key link between content knowledge and support of student learning. Mathematical knowledge for teaching (MKT) has been described as the domains of knowledge that include a teacher's subject matter knowledge and her pedagogical content knowledge (Ball, 1990; Hill, Ball, & Schilling, 2008). MKT has also been described as a teacher's key developmental understandings and how they influence a teacher's practice (Silverman & Thompson, 2008). It has been reported that many teachers do not possess key developmental understandings (KDUs) of central ideas of secondary mathematics, and that these understandings can only emerge from experiences that promote perturbations that result in self-reflection.

A teacher's mathematical teaching orientation influences her classroom practices (A. G. Thompson, Philipp, Thompson, & Boyd, 1994). A teacher with a calculational orientation has an image of mathematics as an application of rules and procedures for finding numerical answers to problems. A teacher having a conceptual orientation has an image of mathematics as a network of ideas and relationships among these ideas, and strives to support students in developing coherent meanings among these ideas.

I will define a teacher's goal as a mental representation of what a teacher is trying to accomplish. This is similar to how other researchers (Locke & Latham, 2002; Norman, 2002; Pintrich, 2000; Schoenfeld, 1998) have categorized goals, although this perspective does not explain possible purposes or reasons why a teacher may pursue a goal. Research has shown that a teacher's goals for student learning do influence her development of powerful pedagogical content knowledge (Webb, 2011). Other studies have shown that a teacher's mathematical knowledge for teaching also influences her pedagogical goals and actions (Marfai & Carlson, 2012; Moore, Teuscher, & Carlson, 2011). The relationship between a teacher's goals and her mathematical knowledge for teaching are reciprocal; each influences the other.

In the study on which this preliminary research report is based, I used Silverman and Thompson's (2008) construct of MKT as a lens for examining how a teacher understands ideas and connections among ideas, and how this influences her pedagogical decisions and actions. The transformation of a teacher's key developmental understandings (Simon, 2006) into MKT is developmental as a teacher's orientation shifts from calculational to conceptual, and I hypothesize examining teachers' pedagogical goals for a lesson can lead to insights underlying this process of growth. Figure 1 illustrates the interactions within this theoretical framework.

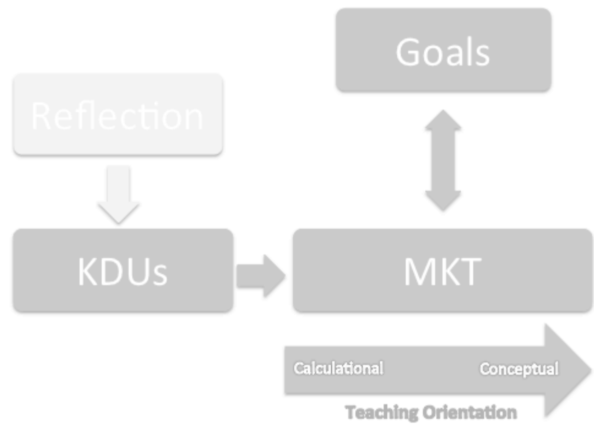


Figure 1. Interactions within the theoretical framework

As illustrated in Figure 1, I make the claim that through the process of self-reflection, a teacher's orientation may shift from calculational to conceptual, and not the other way around. A teacher's mathematical orientation is influenced by her MKT and that this knowledge impacts the goals a teacher has for her students' learning and her teaching. A teacher forms new KDUs as she makes more connections between key ideas of mathematics through the process of self-reflection; these conceptual advances in a teacher's understanding of mathematics are supportive of a conceptual orientation.

Research Question

My research question in this study was as follows. How might a teacher's pedagogical goals for student learning be characterized in the context of using a curriculum promoting a conceptual orientation of mathematics, and how are they similar or different than when using a curriculum that promotes a calculational orientation?

Methods

To characterize a teacher's goals for student learning, I used a goal framework that had emerged from a prior study using grounded theory (Strauss & Corbin, 1990), keeping the perspective of Silverman and Thompson's characterization of MKT and teaching orientation in mind. The goal framework is given in Table 1.

Table 1. Levels in a Teacher's Goals for Student Learning

Goal Coding	Description
TGSL0	Goals are not stated, or the teacher states that the goals of the lesson are unknown.
TGSL1	Goals are a list of topics that a teacher wants her students to learn in the lesson, each associated with an overarching action.
TGSL2	Goals are a list of topics that a teacher wants her students to learn in the lesson, each associated with a specific action.

Goal Coding	Description
TGSL3	Goals are doing methods of mathematics that a teacher wants her students to learn in the lesson.
TGSL4	Goals are getting students to think about the mathematics in the lesson, without the ways of thinking articulated.
TGSL5	Goals are getting students to think about the mathematics in certain ways during the lesson.
*TGSL6**	Goals are about developing ways of thinking about the mathematics in the lesson, with attention to how that thinking may develop.
*TGSL7**	Goals are about developing ways of thinking about the mathematics in the lesson, with attention to how that thinking may develop, and how these ways of thinking may be either mathematically powerful or hindering for future learning.

Goal levels marked with an asterisk (TGSL6 and TGSL7) were not observed in the prior study, but were hypothesized to exist based on researchers' work in professional development supports that promoted growth of teachers' goals attentive to students' thinking of mathematics and ways to support such thinking (Smith, Bill, & Hughes, 2008; Smith & Stein, 2011; P. W. Thompson, 2009).

The context in which the goal framework originally emerged focused on characterizing the participants' pedagogical goals for student learning with a research-based Precalculus curriculum (Carlson & Oehrtman, 2012) from the Pathways project that promoted well-connected understandings of quantities, covariation, proportionality, the constant rate of change, and the average rate of change. Some examples of teachers' goals for student learning from this prior study included "Discuss three ways (ratio, constant multiple, scaling) quantities are proportional" (TGSL3), "State quantities precisely - don't use pronouns (want the object, the attribute of that object, units of measure)" (TSGL2), "Develop the equation of a circle" (TGSL1), and "Get through it" (TGSL0). Project Pathways is an initiative that focuses on professional development to improve teachers' key developmental understandings of the mathematics they teach in order to improve content knowledge through fostering a rich connection of mathematical ideas and relationships, and part of this initiative led to project leaders developing a research-based conceptually oriented curriculum that teachers involved on the project would use in their classroom.

In the current study, Robert (pseudonym), a secondary mathematics teacher from Salt Valley High School (pseudonym) in a Southwestern state was selected for observation during two chapters in which he taught Trigonometry during the Spring 2013 semester. Robert was teaching Precalculus for the third time using the same conceptually rich curriculum as the teachers in the prior study, although he had supplemented the course with materials from a traditional textbook. Robert has been teaching for 13 years total at the same high school. Robert was identified by the Pathways project as a teacher whose key developmental understandings of the Precalculus curriculum were well connected and whose pedagogical actions indicated an inclination to act on student thinking.

To characterize Robert's lesson-specific goals for student learning, twenty-nine classroom observations were videotaped that primarily covered two chapters from different texts focusing on trigonometry, in particular angle measure, trigonometric functions, identities, and applications using trigonometric functions. Field notes were taken during each observation. Prior to the series of classroom observations, an initial questionnaire was given to characterize Robert's overarching goals. At the end of class, the researcher gave a short questionnaire that included queries about Robert's instructional goals and his goals for student learning that day. Robert responded the same day; the norm was established that the researcher would follow up with one or two questions based on his responses to the questionnaire that same day. Analysis in this report specifically focuses on Robert's initial response to the question "What were your goals of instruction with regards to student learning for the lesson you had today?" Robert's goals for student learning were then coded using the goal framework described in Table 1.

For an example of how the coding was done in this study, in a lesson in which the key idea was having students make the connection that a measure of an angle's openness is the quantification of the fraction of any circle's circumference subtended by the angle (with tasks designed to help support students' development of meaning for angle measure in both degrees and radians), Robert's goals for student learning were as follows.

> One goal was for students to gain an understanding of what it means to measure an angle. Another goal was for students to gain an understanding of what it means for an angle measure to be 1 degree. I wanted to stress the importance of thinking about an angle as an object that cuts off a certain fraction of a circle's circumference whose center is the vertex of the angle.

Robert's statement of his goals for student learning that day had two goals followed by one clarifying statement. His first stated goal of what means to measure an angle, which included the clarifying statement, was rated at a TGSL5 level, since the desired way of student thinking was described specifically. However his second stated goal of having students gain an understanding what it means to measure an angle of one degree suggested a desired way of student thinking but was not articulated, and therefore was rated at a TGSL4 level.

The chapters under which the observations were performed had a conceptually rich chapter from the research-based curricular materials, which was then followed by a skill-based chapter and sections from a traditional textbook. In addition to characterizing how curricular context affected Robert's goals for student learning, the researcher also tested the stability of Robert's goals through follow-up questions designed to perturb his goals to higher levels in the framework.

Results

After coding Robert's goals, statements of his goals for student learning ranged from levels 1 to 5 (see Table 2, next page). Based on classroom observations and field notes, in a majority of class sessions Robert made pedagogical moves to model student thinking and he made decisions to act on his model of student thinking either at the group level or in a whole class discussion, with varying levels of success. Robert's pedagogical moves in more successful interactions initially suggested that he was mindful of student thinking and how to support student thinking in the planning process, which implies that goals rated at TGSL6 were accessible to Robert. However, such goals were not stated explicitly.

Goals rated at a TGSL6 level only emerged through the process of follow-up questions, and in a few instances, Robert's responses and his pedagogical moves in the subsequent lesson suggested reflection on how student thinking about the mathematics of the lesson could be promoted or developed. However, these goals never became part of Robert's regularly stated goals for student learning, and therefore were not coded.

Table 2. Teacher's Goals for Student Learning - Count (Percentage)

Goal Level	Conceptual Curriculum	Skill-based Curriculum
TGSL0	0 (0.0%)	0 (0.0%)
TGSL1	7 (17.1%)	2 (9.5%)
TGSL2	10 (24.4%)	1 (4.8%)
TGSL3	3 (7.3%)	8 (38.1%)
TGSL4	17 (41.5%)	7 (33.3%)
TGSL5	4 (9.8%)	3 (14.3%)
TGSL6	0 (0.0%)	0 (0.0%)
All Stated Goals	41 (100.0%)	21 (100.0%)

In looking at the top two categories where Robert's goals for student learning clustered, when using a conceptually rich curriculum the top two ranked categories were TGSL4 (41.5%) and TGSL2 (24.4%), while when using a skill-based curriculum the top two ranked categories were TGSL3 (38.1%) and TGSL4 (33.3%).

Discussion

Although Robert's goals of having students thinking about the mathematics in the lesson (TGSL4) was prominent regardless of the curriculum type used, goals articulating specific methods of mathematics Robert wanted his students to use (TGSL3) topped the types of goals Robert stated for a skill-based curriculum, while goals stating specific actions (TGSL2) in support of mathematical topics were common while using a conceptually rich curriculum. Pintrich (2000) found that "strong" curricular or classroom contexts influenced the types of goals teachers would normally access, so the findings of this study might not be entirely surprising. However, the results may be viewed as surprising if the goal framework is thought of as a trajectory of teacher growth representing a teacher's developing MKT. From this perspective, Robert's goals would seem to show his MKT had more well-connected understandings when using a skill-based curriculum than when using a conceptually rich curriculum, since goals ranked at TGSL3 are higher than TGSL2. However such an analysis would not be appropriate given that we are comparing two different curricula that are supportive of contrasting teaching orientations. A skill-based curriculum and the activities found in it are supportive of a calculational orientation, so goal statements focusing on methods of mathematics a teacher wants her students to do are representative of a different view of the mathematics than goal statements focusing on methods of mathematics viewed from a conceptual orientation. So although the goal framework can be thought of as representing stages of growth of a teacher's goals supportive of student learning, how these goals manifest themselves with teachers having a calculational orientation, versus how these goals manifest themselves with teachers having a conceptual orientation, would need further study.

Although it may seem that the different curricula Robert used promoted different types of goals for student learning, in retrospective analysis of Robert's overarching goals, the results of this study may have been foreshadowed. In an initial questionnaire Robert was asked, "Are these goals [for student learning] affected by the type of lesson you have-for example, a conceptual versus skill based lesson? If yes, how are they affected? If not, how are they not affected?" Robert's response to this question was as follows.

> The over-arching goal of improving student understanding remains for any lesson, regardless of the emphasis of skills vs. concepts. However, the trajectory and/or delivery method of the lesson can be affected. I visualize a

> concept-based lesson as having student investigation as a major portion of the activities, while a skill-based lesson is still focused on "why" certain procedures are done but there is more direct instruction of those procedures.

Looking at Robert's overarching goals, improving student understanding was consistent with goals that promoted students' thinking about the mathematics in the lesson (TGSL4). Robert's mention of student investigations as being a large part of conceptually based lessons promoted goals concerning specific actions he wanted students to take during the course of these investigations (TGSL2). In the last part of his response, Robert's mention of direct instruction of procedures during a skill-based lesson suggested goals focused on methods of mathematics (TGSL3). Looking at Table 2, Robert's overarching goals were predictive of the clustering of top ranked goals he had for student learning for the lessons under which the classroom observations were made.

Although the goal framework contained goals rated at TGSL6 and TGSL7 based on other researchers' findings, these goals did not emerge as a response to the initial question "What were your goals of instruction with regards to student learning for the lesson you had today?" Goals rated at TGSL6 emerged through follow-up questions to stimulate reflection, or were evident based on Robert's pedagogical moves made to support the development of his students' ways of thinking about mathematics. However, the follow-up questions did not perturb Robert's responses to the initial question at any time during this study. Part of this may have to do with the way the question was interpreted by Robert. In a post-study interview, Robert was asked how he interpreted the initial question. His response was that he interpreted the question was asking what he wanted students to do or to understand. By the nature of this interpretation, it eliminated the possibility that Robert would state goals beyond a level ranked at TGSL5. However such an interpretation of goals for student learning is not unique to Robert.

Many school districts in Arizona, the United States, and Canada have incorporated rubrics to assess a teacher's goals for student learning as part of the protocol used in their evaluation. For example, teachers must write or post each day's learning goals on the wall or board so that students (and evaluators who visit their classroom) can clearly see it. For example, in guidelines for assessment, evaluation, and reporting by the Ministry of Education in Ontario, Canada, the section regarding assessment for learning states "learning goals clearly identify what students are expected to know and be able to do, in language that students can readily understand." (Ontario Ministry of Education, 2010, p. 33). Similar statements regarding learning goals can be found in evaluation rubrics for school districts in the United States. It is not surprising that a teacher's stated goals for student learning do not mention ways to promote or support student thinking; there are strong societal norms that push back against such an interpretation. It may be a contributing factor to why Robert's stated goals for student learning did not shift after attempts were made to perturb him toward stated goals that would be at higher levels in the framework.

Conclusion

There are several findings that resulted from this study which inform the direction my future studies will take. First, a teacher's overarching goals appear to be predictive of their lesson specific goals. It suggests that sustainable professional development efforts to perturb goals at the lesson level should also include moves to perturb a teacher's overarching goals that value attention to how a student's thinking about mathematics may be promoted or developed. The second finding about how the phrase "goals for student learning" is interpreted almost universally in United States and Canada suggests a specific meaning is attached to this phrase that is not easily perturbed. Since societal norms push against a broader interpretation of goals for student learning, this suggests efforts to perturb goals toward TGSL6 and TGSL7 may be more successful through directed professional development efforts, such as professional learning communities or through adapting existing protocols, such as the Thinking Through a Lesson Protocol (Smith, et al., 2008) or the Professional Development Spiral (P. W. Thompson, 2009). The adapted protocol would be designed to foster growth of both lesson specific and overarching goals for student learning that attend to students' thinking about mathematics and ways to promote and support such thinking, with attention to how such ways of thinking may help or hinder future learning for the student.

Acknowledgement

Research reported in this paper was supported by National Science Foundation Grant No. EHR-1050721. Any conclusions or recommendations stated here are those of the author and do not necessarily reflect official positions of NSF.

References

Ball, D. L. (1990). The Mathematical Understandings That Prospective Teachers Bring to Teacher Education. *The Elementary School Journal, 90*(4), 449-466.

Carlson, M. P., & Oehrtman, M. C. (2012). *Precalculus: Pathways to Calculus* (3rd ed.). Gilbert, AZ: Rational Reasoning.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education, 39*(4), 372-400.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey. *American Psychologist, 57*(9), 705–717.

Ma, L. (1999). *Knowing and Teaching Elementary Mathematics*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Marfai, F. S., & Carlson, M. P. (2012). *The Influence Of A Teacher's Orientation On Her Mathematical Knowledge For Teaching And Goals.* In L. R. Van Zoest, J.-J. Lo & J. L. Kratky (Eds.), Proceedings of the Thirty-fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Kalamazoo, MI: Western Michigan University. Available at http://pathways.mspnet.org/index.cfm/showcase.

Moore, K. C., Teuscher, D., & Carlson, M. P. (2011). *Exploring Shifts In A Teacher's Key Developmental Understandings And Pedagogical Actions.* In T. d. S. Lamberg & L. R. Wiest (Eds.), Proceedings of the Thirty-third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Nevada: University of Nevada, Reno.

Norman, D. A. (2002). *The Design of Everyday Things*. New York: Basic Books.

Ontario Ministry of Education. (2010). *Growing Success. Assessment, Evaluation, and Reporting in Ontario's Schools*. Retrieved from http://www.edu.gov.on.ca/eng/policyfunding/success.html.

Pintrich, P. R. (2000). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology, 25*(1), 92-104.

Schoenfeld, A. H. (1998). Toward a theory of teaching-in-context. *Issues in Education, 4*(1), 1-94.

Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education, 11*, 499-511.

Simon, M. A. (2006). Key Developmental Understandings in Mathematics: A Direction for Investigating and Establishing Learning Goals. *Mathematical Thinking and Learning, 8*(4), 359-371.

Smith, M. S., Bill, V., & Hughes, E. K. (2008). Thinking Through a Lesson: Successfully Implementing High-Level Tasks. *Mathematics Teaching In The Middle School, 14*(3), 132-138.

Smith, M. S., & Stein, M. K. (2011). *5 Practices For Orchestrating Productive Mathematics Discussions*. Thousand Oaks, CA: NCTM/Corwin.

Sowder, J. T. (2007). The Mathematical Education and Development of Teachers. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 157-224). Charlote, NC: Information Age Publishing.

Stigler, J. W., & Hiebert, J. (1999). *The Teaching Gap*. New York: Free Press.

Strauss, A. C., & Corbin, J. M. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*: Sage Publications.

Thompson, A. G., Philipp, R. A., Thompson, P. W., & Boyd, B. A. (1994). Calculational and Conceptual Orientations in Teaching Mathematics. In A. F. Coxford (Ed.), *1994 Yearbook of the NCTM* (pp. 79-92). Reston, VA: NCTM.

Thompson, P. W. (2009). *Case Modules Facilitator Guide*.

Webb, M. (2011). *Teachers' Knowledge Is Built Around Their Goals.* In T. d. S. Lamberg & L. R. Wiest (Eds.), Proceedings of the Thirty-third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Nevada: University of Nevada, Reno.

17

MATHEMATICS BEYOND THE CLASSROOM: STUDENTS' 'VALUE CREATION' THROUGH MATHEMATICAL MODELING WITHIN A LEARNING COMMUNITY

Joo young Park
Teachers College, Columbia University

This study examined how mathematical modeling activities within a collaborative group impact on students' perceived 'value' of mathematics. With a unified framework of Makiguchi's theory of 'value', mathematical disposition, and identity, the study identified the elements of the value-beauty, gains, and social good-with the observable evidences of mathematical disposition and identity. A total of 60 college students participated in 'Lifestyle' mathematical modeling project. Both qualitative and quantitative methods were used for data collection and analysis. The result from a paired-samples t-test showed the significant changes in students' mathematical disposition. The results from the analysis of students' written responses and interview data described how the context of the modeling tasks and the collaborative group interplayed with students' perceived value.

Key words: Mathematical Modeling, Instructional Activities and Practice, Value Creation, Mathematics Disposition

Introduction

Studies reported that when students see themselves as capable of doing well in mathematics, they tend to value mathematics more than students who do not see themselves as capable of doing well (Eccles, Wigfield, & Reuman, 1987; Midgley, Feldlaufer, & Eccles, 1989). To see the value in mathematics, it is essential for students to believe that mathematics is understandable, not arbitrary; that, with diligent effort, it can be learned and used; and they are capable of figuring out mathematical problems based on their experiences. Kilpatrick and his colleagues (2001) introduced "productive disposition" as one of key components of mathematical proficiency and defined as the "habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy"(NRC, 2001, p. 131). "Mathematics disposition" appeared in the National Council of Teachers of Mathematics Evaluation Standards as "a tendency to think and to act in positive ways"(NCTM, 1989, p233), which is manifested when students approach tasks.

Developing such a disposition toward mathematics requires frequent opportunities to recognize the benefits of perseverance and to experience the rewards of sense making in mathematics. It becomes a question of what learning environment supports students to engage in meaningful learning of mathematics and to develop positive disposition as well as self-concept. A number of studies demonstrated that mathematical modeling, which plays a prominent role in the new Common Core State Standards for Mathematics (CCSSM), promotes socially situated learning environments with group collaboration, classroom discussion, initiative, and creativity, and it has the potential to develops positive disposition toward mathematics and strengthen their mathematical identity (Ernest, 2002; Lesh & Doerr, 2003). The studies highlight that learning mathematics extends beyond individuals' learning concepts, procedures, and learners learn to be a part of a community of practice and become participants in the mathematics being practiced (Boaler, 2002).

Theoretical Framework

As a unified framework of Makiguchi's theory of value creation (1930), mathematical dispositions outlined by NCTM Evaluation Standards 10, and identity (see Table 1), this

study identified the elements of the value with the observable evidences of mathematical disposition and identity. The concept of value in the notion of Makiguchi (1930; Bethel, 1989) takes into account the subject and object relationship (students' relationship with mathematics in this study), which reflects human creativity. In the notion of Makiguchi (1930; Bethel, 1989)'s value creation, it is critical that students feel happiness, enjoyment, and pleasure in their own processes of investigating and understanding mathematics, as a result, students construct meaning, and value is created. In Makiguchi's concept of value, the three elements of the value are the following:
Beauty is perceived to be an emotional and temporary value. The value of *Gain* is an individual value and self-development, and beneficial aspect that is related to the whole of man's life. *Social good,* however, is a social value and is related to the life of the group. The value of good is the expression given to the evaluation of each individual's voluntary action, which contributes to the growth of a unified community composed of the individuals (Makiguchi, 1930; Bethel, 1989).

Makiguchi's Theory of Education

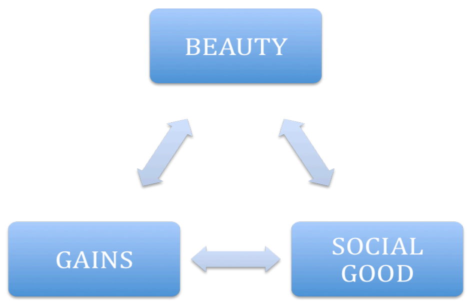


Figure 1- Makiguchi's theory of value creation (Makiguchi, 1930; Bethel, 1989)

Based on Makiguchi's philosophy, what is important to create value is to pursue gain, good, and beauty- without overemphasizing one or ignoring the others in a balanced and mutually reinforcing manner.

Table 1. Theoretical framework (Makiguchi's theory of value, disposition, and identity)

Mathematical Disposition, Identity, Sense of belonging	Makiguchi's Elements of Value
• Interest, curiosity, and inventiveness in doing math	Beauty
• Confidence in using math to solve problems and communicate ideas • Willingness to persevere and become persistent in math tasks • Flexibility in exploring math ideas and trying alternative methods in solving problems • Appreciation of the role of mathematics in our culture and its value as a tool and as a language • Inclination to monitor and reflect on their own thinking and performance • Valuing of the application of mathematics to situations arising in other disciplines and everyday experiences	Gains
• See oneself as a learner, and doer of mathematics (Identity) • Sense of belonging in a learning community, global citizenship	Social value (Social Good)

The theory of value creation (Makiguchi, 1930; Bethel, 1989) shares a common thread with a mathematical modeling perspective. *Mathematical modeling* is the process of leading from a problem situation to a mathematical model (Kaiser, 2006). A number of studies reported that mathematical modeling promotes socially situated learning environments with group collaboration, classroom discussion, initiative, and creativity and it has the potential to empower students and strengthen their mathematical identity (Ernest, 2002; Lesh & Doerr, 2003). These studies highlight that learning mathematics extends beyond individuals' learning concepts, procedures, and learners learn to be a part of a community of practice and to become participants in the mathematics being practiced (Boaler, et al., 2000). The modeling aspect shares the common threads with the view of Makiguchi (1930), which emphasized the engagement of students in the actual activities of the community and urged the students to engage with the challenges of problem solving and finding creative responses to the problems we find in our everyday lives. How a student learns mathematics involves the development of the student's identity as a part of a mathematics classroom community (Anderson, 2007).

Accordingly, students need an opportunity of sharing meanings and values through engaging in real life mathematics. The assessment of mathematical knowledge needs to include evaluations of these indicators and students' appreciation of the role and value of mathematics. Interaction between individuals and environments (e.g., teachers, peers, curriculum and so on) can be understood in terms of Makiguchi's notion of *value*, specifically, *beauty, gain*, and *social value* derived from how individuals relate to their environment, in this study, the mathematics classroom via mathematical modeling activities. The examination of mathematics disposition gives insights into meaningful integration of cognition and affective skills in learning mathematics.

Research Questions

The purpose of this study is to evaluate an instructional model for students to *create value* in learning mathematics. With the unified framework of Makiguchi's theory of value, mathematics disposition, and identity, this study examines how 'socially-situated' mathematical modeling activity within a collaborative learning community can contribute to students' development of their mathematical disposition, identity, and sense of community as well as students' creating mathematical meaning. The guiding questions for this study are as follows:

1. What changes (if any) are observed in students' mathematical disposition that results from learning mathematics through mathematical modeling within a learning community? Specifically, How do students perceive value of beauty and gains, in Makiguchi's notion, of learning mathematics after experiencing mathematical modeling activities within a collaborative group?

2. How are students' mathematical identities transformed from their involvement in mathematical modeling activities within a collaborative group?

3. How are students' perceived social values, in Makiguchi's notion, of learning mathematics observed during mathematical modeling activities within a collaborative group?

How does the collaborative group create a sense of belonging to the group that can be realized through engaging in mathematical modeling activities with group members?

Methods

Both quantitative and qualitative methodologies were used in data collection and analysis, investigation, and interpretation. Multiple data sources including surveys, interview data, students' written tasks and journals were collected (see figure 1). These data sources provided participants with multiple opportunities of their reflecting and sharing thoughts about how these experiences impacted their disposition and identity. The participants were a total 60 students who enrolled in college algebra courses taught by the researcher. The curricular task for the study was a modified version of the mathematical modeling project developed by the Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) at Rutgers University. The project introduces the ecology of humans as a topic, and ecological foot printing is developed as a tool for assessing human impact and as a decision-making tool. These topics are relevant to social and environmental issues in which students engage in everyday lives. The investigator attempted to provide students with the tasks that require everyday knowledge, critical thinking, and a collaborative work. The mathematical modeling project was conducted within groups of four or five for four consecutive weeks. After completing the first week of conducting the project, students were asked to collect their own data. The Mathematical Disposition Survey (MDS) was conducted at the beginning of the study and the end of the study, and the results were compared. The mathematical disposition survey instrument is a modification of the one developed by Kisunsu (2008). Students' written tasks and journals were collected after each class. A total of eighteen focal students were selected for interview based the results from the analysis of Mathematical Disposition

Survey and students' journals. Semi-structured interview offered students the opportunity of giving detailed statements on their written tasks, questionnaires, and journals. The researcher took field notes and audio-taped all the activities in classroom and interviews.

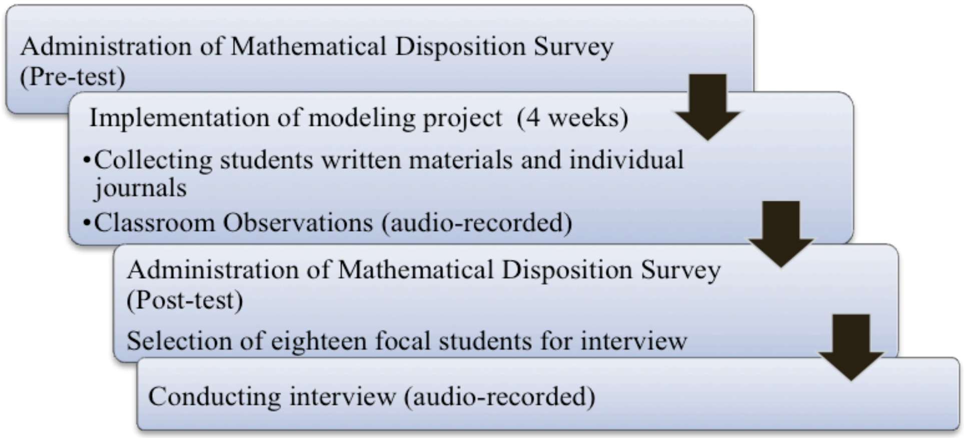


Figure 2- Procedures (Multiple methods)

The coding scheme was created in terms of the components of disposition and values when finding common themes and aiming to answer each research question. To validate whether different coders would code the same data the same way, the two responsible for coding transcripts were the investigator and a professional researcher in mathematics education. Two separate coding documents were created for coders. One provided the list of codes with examples of utterances associated with each code and guidelines for coding lesson transcripts. The other provided practice in coding and was used in the training sessions. Initially, coders worked completely separately. However, the coders first coded independently and then met to try to reach a consensus on their coding transcripts. Intercoder agreement for each of the thematic categories was calculated, as well as an overall average agreement across all the themes. There is also quantitative data analysis in this study. Matched pair t-distributions were used to determine the overall outcome of the Disposition Inventory and outcome in each component of Disposition (Appreciation, Interest, Usefulness, Persistence, Meta-cognition, Flexibility, Confidence, and Modeling) Inventory.

Results

The result from a paired samples t-test showed the significant changes in students' mathematical disposition between pre and post survey. There was significant difference in the mean scores for Mathematics Disposition Pre-test (Mean =132.57, SD= 23.65) and Post-test (Mean=138.97, SD= 24.52 with t=-3.25, $p < 0.01$) (See Table 2.)

Table 2. Descriptive Statistics and Paired Samples T-test (Mathematics Disposition)

	Pre-Test	Post-Test
Mean	132.57	138.97
Variance	567.12	607.65
Observations (N)	47	47
Pearson Correlation	0.85	
Df	46	
t Stat	-3.248	
P(T<=t) two-tail	0.002	

t Critical two-tail	2.013

For further investigation, the pre and post survey mean scores in each aspect of mathematics disposition (confidence, flexibility, perseverance, interest and inventiveness, meta-cognition, usefulness and appreciation) were analyzed by a paired samples t-test. Students' gain score from pre to post test was statistically significant for the aspect of flexibility (Pre: Mean=10.53, SD=3.69; post: Mean=12.09, SD=2.51 with t=-3.28, $p<0.01$), for the aspect of appreciation (Pre: Mean=16.49, SD=5.34; Post: Mean=18.06, SD= 4.61 with t=-2.62, $p<0.01$). Interview data were analyzed searching for evidence of 'changes' in disposition, identity, and students' perceived value resulting from engaging in modeling activities in a collaborative group.

Students' changes in disposition and identity

Interview data were analyzed searching for evidence of 'changes' in disposition resulting from engaging in modeling activities in a collaborative group. Two coders divided the number of coding agreement by the number of agreement and disagreements combined. For instance, with two coders, if 18 text units had been coded "change in the aspect of appreciation" by at least one of two coders and in 15 of those cases both had invoked the code on the same text, then the level of intercoder reliability (Krippendorff, 2004; Campbell et al, 2013), would be approximately 83% for students' change in the aspect of appreciation. I calculated overall intercoder reliability for all codes as a set by diving the total number of agreements for all codes by the total number of agreements and disagreements for all codes combined (Campbell et al, 2013). The overall intercoder reliability was .82.

When analyzing interview data from seventeen participants, I discovered that 82% of the interview participants reported disposition changes in the aspect of appreciation of disposition, 58% of the students reported their changes in the aspect of usefulness, and 70% of the participants reported changes in the aspect of modeling. Most participants reported changes in more than one aspect. The results were nearly consistent with the statistical analysis in that a higher percentage of participants changed disposition in the aspect of appreciation, modeling, and usefulness than in the other aspects.

One common characteristic emerged from interview data was that students attributed the changes in their own mathematics dispositions to their appreciation of the real life context of the modeling project:

Excerpt 1. Change in the aspect of usefulness and appreciation

> Chloe: Since I do not use math much, I did not recognize the importance of math before. It was the first experience of seeing how much math was involved in everyday life. I didn't see the importance of math as much as I do after this project.
>
> Lisa: I think it's important to know since it's environmental issue and I thought it was interesting topic. I never did like this before. It was interesting to see how much we used.

The aspect of modeling that was relevant to students' everyday lives seemed to have contributed to their development of positive disposition and personal identity as doers of mathematics:

Excerpt 2: Change in the aspect of interest and confidence, and identity transformation

Ella: I had difficult in doing math entire my life until to this day. I feel like this project would be beneficial for the students like me. I think it is important to think analytically and think outside of box through this kind of project. I changed my view of math in the sense that it became enjoyable since this project gave me some excitement. The project used math but it was interesting.

Excerpt 3. Change in the aspect of interest and confidence

Jessica: Usually, sitting and doing are not for fun but this project made me to enjoy the process of doing it and liked it. Since it was an ongoing thing, I did not want to miss the class cause if I missed a part of the project, I know that I could not catch up nor finish it.

Jade: I became more confident and suggested how we should solve the problem. So I think that this project was pretty interesting cause I usually tend to lay back, and everyone else takes control but for this project, I took control, gave opinions of how we should find the solutions. It was a learning experience that actually I used my brain to answer the questions.[…] I think the best moment was when I got the solution which was the same as everyone else's. I had never had that experience. That was the best moment.

Students came to value and appreciate the modeling activities of creating and interpreting data in real life context, and developed their autonomous modeling behavior:

Herald: By breaking it down, it was easier to see the steps, instead of using a given equation. […] So I think it is much easier to do steps by yourself instead someone else is telling you what to do or how to do.

Aubrey: I think again, the questions like, the conversions, how to convert one unit to the other, then a question like "where does the excess land come from to support people in US?", we had to think about what is the bio capacity of US, it involves critical thinking by having compared numbers and looked back our data. It was great. We had never done like this before. We usually just did normal math procedures in our math class.

As the above excerpts indicate, students seem to have developed identities as doers of meaningful mathematics; students expressed themselves to be people who want to use their own ideas, exercise their own thought, and think critically.

Sense of belonging and social value: what it means to understand mathematics

There was evidence indicating socialization in a group through emotional connections by asking for help and sharing stories of events with particular topics:

Lisa: We talked about our data, also our personal lives, why we had these numbers, and what electric devices have used. I had a big number [footprint] but she had a smaller number than mine. Then we talked about why and talked about the details in our personal lives. Especially with Alexis, cause we both live with family but others live on campus so our numbers were pretty close but others' were very different from ours. We talked about it at the personal level. Generally, as for a group work, some people do not do their parts but in our case, everyone contributed their parts.

Interpreting mathematical results and social value: what counts as mathematical argumentations in modeling tasks

‘What constitutes mathematics argument’ was related to the affiliation with the group, and the real life context of modeling tasks helped them to establish socio- mathematical norms.

Interviewer: While working on this project in a group, how did you all validate that your solution is correct?

> *Deana*: We took a look at them to see if they make sense, like realistic numbers not too high or too low. One girl’s number was so low and everybody else was high, so we told her “you did something wrong”. Then we found that she forgot to add something.

In order to claim if a solution is correct, students seemed to have established socio mathematical norms based on the realistic context of modeling: “if they make sense, like realistic number not too high or too low”. It involves knowing how to examine the data to see if it makes sense in real life situation. This data analysis demonstrates what counts as an acceptable mathematical explanation and mathematical understanding.

Students seem to be able to make a connection not merely with their immediate community (a group in the classroom) but with a global community through interpreting mathematical results. Students also developed their identity as a global citizen:

> Lily: I realized what myself and my family can do for our environment through this project. Small things like, even having one more person in your car really makes a difference. Through calculating the footprint, I found that, you could reduce your footprint by walking, instead of taking a cab. It greatly changes your footprint. When I go to Whole Foods [market] I always use a reusable bag.

As Lily demonstrated, for example, students not only developed their productive relations with mathematics but also their identity as a global citizen.

The modeling task with students’ personal data and social interactions within a group contributed substantially to students’ creating social value. While working with ‘realistic data’, students were able to develop modeling competence: to see the possibilities that mathematics offer for the solution of real world problems and to value them positively (Maass, 2006); to think about the nature of mathematics and assess their own capabilities beyond constructing and investigating mathematical models, and students were aware of the limitations in real life problem situations on validating their data.

Conclusion

In Makiguchi’s theory of value, benefit or gain is a beneficial aspect of the interactions with an object (mathematics in this study). For example, students did develop a greater level of confidence in doing mathematics and were able to express their ideas within a group or developed one’s willingness to navigate alternative ways of solving problems and monitor thinking. The individual creates value through contributing to the well-being of the larger human community and society (Ikeda, 2001). Social value seemed to have been created through students’ interactions with the external context to mathematics and also with other members while working in a group. Students deeply engaged with mathematics through modeling activities by sharing mistakes, listening to and offering suggestions about other’s work, and thinking about rationales behind why particular decisions were meaningful. The development of disposition seems to be shaped by the interrelation between the context of mathematics tasks and interactions with others. For further study, by examining students’ modeling activities and interactions with peers in the classroom, one can understand better how these elements interplay with students' construction of disposition and identity. The underlying assumption of Makiguchi's value creation is that every student has the ability to contribute to his or her own development and that of society in a creative way. Overall, to create value means to self-actualize one's full potential and create beauty, gain, and good, from all circumstances.

The instructional model evaluated in this study-an interdisciplinary mathematical modeling project conducted within a collaborative group-can be implemented in a college or high school classroom as a model that helps students to develop their positive mathematics disposition and critical thinking, to engage in sense-making mathematics, and to create a learning community.

References

Anderson, R. (2007). Being a mathematics learner: four faces of identity. *The Mathematics Educator, 17*(1), 7-14.

Bethel, D. (1989). (Ed.) Education for creative living: Ideas and proposals of Tsunesaburo Makiguchi, Iowa State University Press. An edit and translation into English of Makiguchi(1930)'s work on pedagogy.

Boaler, J. (2002). The development of disciplinary relationships: knowledge, practice, and identity in mathematics classrooms. *For the Learning of Mathematics, 22*(1), 42-47.

Campbell, L., Quincy, C., Osserman, J., & Pedersen, O. (2013). Coding in-depth semi-structured interviews: problems of unitization and inter-coder. reliability and agreement, *Sociologic*.

i.

Cobb, P., Gresalfi, M., & Hodge, L. L. (2008). An interpretive scheme for analyzing the identities that students develop in mathematics classrooms. *Journal for Research in Mathematics Education, 40*(1), pps. 40-68.

Eccles, J., Wigfield, A., & Reuman, D. (1987). *Changes in self-perceptions and values at early adolescence*. Paper presented at the meeting of the American Educational Research Association, Washington, DC.

Ernest, P. (2002). Empowerment in mathematics education. *Philosophy of Mathematics Education Journal* (15).

Ikeda, D. (2001). John Dewy and Tsuneosaburo Makiguchi: Confluences of thought and actions. (Ed.), *Soka Education: For the Happiness of the Individual*. Santa Monica: Middleway Press.

Kisunsu, P. (2008) Teacher instructional practices, student mathematical dispositions, and mathematics achievement. Doctoral dissertation, Northern Illinois. Retrieved from http://www.lib.umi.com/dissertations:ProQuest Digital Dissertations.

Krippendorff K (2004) *Content analysis: an introduction to its methodology*. Sage, Thousand Oaks, Calif.

Lesh, R., & Doerr, H. (2003). Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp.3-34). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Maass, K. (2006) What are modelling competencies? ZDM 38(2), pp 113–142.

Midgley, C., Feldlaufer, H., Eccles, J. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. APA, 81(2), 247-258.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.

National Research Council (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: The National Academies Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge, England: Cambridge University Press.

ARE STUDENTS BETTER AT VALIDATION AFTER AN INQUIRY-BASED TRANSITION-TO-PROOF COURSE?

Annie Selden
New Mexico State University

John Selden
New Mexico State University

*We present the results of a study of the observed proof validation abilities and behaviors of sixteen undergraduates after taking an inquiry-based transition-to-proof course. Students were interviewed individually towards the end of the course using the same protocol that we had used earlier at the beginning of a similar course (Selden and Selden, 2003). Results include a description of the students' observed validation behaviors, a description of their proffered evaluative comments, and the, perhaps counterintuitive, suggestion that taking an inquiry-based transition-to-proof course does **not** seem to enhance validation abilities. We also discuss distinctions between proof validation, proof comprehension, proof construction and proof evaluation and the need for research on their interrelations*.

Key words: Transition-to-proof, Proof, Validation

We present the results of a study of the observed proof validation behaviors of 16 undergraduates after taking an inquiry-based transition-to-proof course emphasizing proof construction. Students were interviewed individually towards the end of the course employing the same protocol used in our earlier study (Selden & Selden, 2003). Here, as in our earlier study, we regard proofs as texts that establish the truth of theorems and use our previous description of *proof validation* as the reading of, and reflection on, proofs to determine their correctness. Indeed,

> A validation is often much longer and more complex than the written proof and may be difficult to observe because not all of it is conscious. Moreover, even its conscious part may be conducted silently using inner speech and vision. Validation can include asking and answering questions, assenting to claims, constructing subproofs, remembering or finding and interpreting other theorems and definitions, complying with instructions (e.g., to consider or name something), and conscious (but probably nonverbal) feelings of rightness or wrongness. Proof validation can also include the production of a new text—a validator-constructed modification of the written argument—that might include additional calculations, expansions of definitions, or constructions of subproofs. Towards the end of a validation, in an effort to capture the essence of the argument in a single train-of-thought, contractions of the argument might be undertaken. (p. 5).

In this paper, we provide detailed descriptions of the observed validation behaviors that our 16 undergraduates took – something either not done, or only partially done, in prior validations studies and perhaps not at all for this level of student. Past validation studies include: first-year Irish undergraduates' validations and evaluations (Pfeiffer, 2011); U.S. undergraduates' validations at the beginning of a transition-to-proof course (Selden & Selden, 2003); U.S. mathematics majors' validation practices across several content domains (Ko & Knuth, 2013); U.S. mathematicians' validations (Weber, 2008); and U.K. novices' and experts' reading of proofs, using eye-tracking, to compare their validation behaviors (Inglis & Alcock, 2012).

Our ultimate goal is to understand both the process of proof construction and the process of proof validation. Our specific research question was: Would taking an inquiry-based transition-to-proof course that emphasized proof construction significantly enhance students' proof validation abilities?

Theoretical Perspective

We view proof construction as a sequence of mental or physical actions in response to situations in a partly completed proof. This process, even when done with few errors or redundancies, contains many more actions, or steps, than appear in the final written proof and cannot be fully reconstructed from a final written proof. Many of these actions, such as "unpacking" the conclusion to see what one is being asked to prove, or drawing a diagram, do not appear in the final written proof, and hence, are unavailable to students for their later consideration and reflection.

Many proving actions appear to be the result of the enactment of small, automated situation-action pairs that we have termed *behavioral schemas* (Selden, McKee, & Selden, 2010; Selden & Selden, 2008). A common beneficial behavioral schema consists of a situation where one has to prove a universally quantified statement like, "For all real numbers x, $P(x)$" and the action is writing into the proof something like, "Let x be a real number," meaning x is arbitrary but fixed. Focusing on such behavioral schemas, that is, on small habits of mind for proving, has two advantages. First, the uses and interactions of behavioral schemas are relatively easy to examine. Second, this perspective is not only explanatory but also suggests concrete teaching actions, such as the use of practice to encourage the formation of beneficial schemas and the elimination of detrimental ones. (See the case of Sofia and her "unreflective guess" behavioral schema in Selden, McKee, and Selden, 2010, pp. 211-212).

While we have investigated, and written about, a number of proof construction actions (e.g., Selden, McKee, & Selden, 2010), our thinking about proof validation actions is still in its infancy. However, it seems reasonable to conjecture, based on the extant proof validation literature (e.g., Inglis & Alcock, 2012; Selden & Selden, 2003, Weber, 2008) that examination of the overall structure of a proof is crucial in order to determine whether the given attempted proof, if correct, actually proves the statement (theorem) that it sets out to prove. In addition, it also seems that a careful line-by-line reading of an attempted proof is useful for determining whether the individual assertions are warranted, either explicitly or implicitly (e.g., Weber & Alcock, 2005).

Setting of the Research: The Course and the Students

The course that the participants attended is meant as a second-year university transition-to-proof course for mathematics and secondary education mathematics majors, but is often taken by a variety of majors and by more advanced undergraduate students.[1] The course was given at a Southwestern Ph.D.-granting university and was taught in a very modified Moore Method way (Coppin, Mahavier, May, & Parker, 2009; Mahavier,1999). That is, students were given course notes with definitions, questions, requests for examples, and statements of theorems to prove. In addition, the course notes contained one very detailed sample set theory proof construction, some explanations of types of proof frameworks (Selden & Selden, 2009, 1995) and a number of operable interpretations of definitions.[2] That is, we included

[1] We have found that students are often afraid of a transition-to-proof course, and that instead of taking it in their second year of university, before courses like abstract algebra and real analysis (with which it is supposed to help), they take it later.

[2] Bills & Tall (1998, p. 104) considered a definition to be *formally operable* for a student if that student "is able

statements about how to use a definition in proving a theorem and statements about how to prove that a definition is satisfied. We have found this level of detail useful, sometimes even necessary, for our second-year university students.

To illustrate this, we have found that the formal definition of $f(A) = \{ y \mid \text{there is an } a \in A \text{ so that } f(a) = y \}$ is difficult for students to use. So we have added a link to the following *operable interpretation*: To *show* "$y \in f(A)$" you show "there is an $a \in A$ such that $f(a) = y$." To *use* "$y \in f(A)$" you may say "there is an $a \in A$ such that $f(a) = y$." Yet, despite having included such operable interpretations, we have sometimes overheard students, during group work in a subsequent similar transition-to-proof course, say of our operable interpretations that they don't know what these mean. It is now our conjecture that students may, in addition, need examples of how to use a definition in proof construction and of how to show that a definition is satisfied.

The students in this study proved the theorems outside of class and presented their proofs in class on the blackboard and received extensive critiques. These critiques consisted of careful line-by-line readings and validations of the students' proof attempts, often with corrections and insertions of missing warrants. In a sense, the second author modeled proof validation for the students. This was followed by a second reading of the students' proof attempts, indicating how these might have been written in "better style" to conform to the genre of proofs (Selden & Selden, 2013). Once these corrections and suggestions had been made, the student, who had made the proof attempt, was asked to write his/her proof up carefully, including the corrections and suggestions, for duplication for the entire class. In this way, by the end of the semester, the students had obtained one correct, well-written proof for each theorem in the course notes. In addition, about once a week, the class worked in groups to co-construct[3] proofs of upcoming theorems in the course notes. Sometimes, if the students seemed to need it, there were mini-lectures on topics such as logic or proof by contradiction. These mini-lectures were not preplanned; rather they occurred spontaneously, as the need arose.

The homework, assigned each class period, consisted of requests for proofs of the next two or three theorems in the course notes. These proof attempts were handed in at the beginning of the next class to the first author, who determined "on the spot", based on the students' written work, which students would be asked to present their proof attempts on the blackboard that day. The students were aware that being asked to present their proof attempts did not necessarily mean that these were correct, but rather that their proof attempts would probably provide interesting points for the second author to discuss. In addition to presenting their attempted proofs in class, the students had both mid-term and final examinations, which consisted of theorems, new to them, to prove. The mathematical topics considered in the course included sets, functions, continuity, and beginning abstract algebra in the form of a few theorems about semigroups and homomorphisms. However, the teaching aim was to have students experience constructing as many different kinds of proofs as possible, especially in abstract algebra and real analysis, and *not* to learn a particular mathematical content. The course notes were self-contained, that is, all relevant definitions were provided.

Methodology of the Study: The Conduct of the Interviews

Sixteen of the 17 students enrolled in the course opted to participate in the study for extra credit. Of these, 81% (13 of 16) were either mathematics majors, secondary education

---

to use it in creating or meaningfully reproducing a formal argument [proof]".

[3] A detailed description of the co-construction of real analysis proofs is given in McKee, Savic, Selden, and Selden (2010)

mathematics majors, or were in mathematics-related fields (e.g., electrical engineering, civil engineering, computer science).

Interviews were conducted outside of class during the final two weeks of the course. The students received extra credit for participating and signed up for convenient one-hour time slots. They were told that they need not study for this extra credit session. The protocol was the same as that of our earlier validation study (Selden & Selden, 2003) and is reproduced in Appendix 2.

Upon arrival, participants[4] were first informed that they were going to validate four student-constructed "proofs" of a single number theory theorem, indeed, that the proof attempts that they were about to read were submitted for credit by students, who like themselves, had been in a transition-to-proof course. The participants were asked to think aloud and to decide whether the purported proofs were indeed proofs. Participants were encouraged to ask clarification questions and informed that the interviewer would decide whether a question could be answered. They were given the same Fact Sheet (Appendix 1) about multiples of 3 provided to the participants of our earlier study (Selden & Selden, 2003).

There were four phases to the interview: A warm-up phase during which the participants gave examples of the theorem: *For any positive integer n, if* n^2 *is a multiple of* 3, *then n is a multiple of* 3 and then tried to prove it; a second phase during which they validated, one-by-one, the four purported (student-constructed) proofs of the theorem; a third phase during which they were able to reconsider the four purported proofs (presented altogether on one sheet of paper), and a fourth debrief phase during which they answered questions about how they normally read proofs. (See Appendix 2 for details.)

The interviews were audio recorded. The participants wrote as much or as little as they wanted on the sheets with the purported proofs. Participants took as much time as they wanted to validate each proof, with one participant initially taking 25 minutes to validate "Proof (a)".

The interviewer answered an occasional clarification question, such as the meaning of the vertical bar in $3|n^2$, but otherwise only took notes, and handed the participants the next printed page when they were ready for it.

The data collected included: the sheets on which the participants wrote, the interviewer's notes, and the recordings of the interviews. These data were analyzed multiple times to note anything that might be of interest. Tallies were made of such things as: the number of correct judgments made by each participant individually; the percentage of correct judgments made by the participants (as a group) at the end of Phase 2 and again at the end of Phase 3; the validation behaviors that the participants were observed by the interviewer to have taken; the validation comments that the participants proffered; the amount of time taken by each participant to validate each of the purported proofs; the number of times each participant reread each purported proof; the number of participants who underlined or circled parts of the purported proofs; the number of times the participants substituted numbers for n; and the number of times the participants consulted the Fact Sheet. Many of these are indicated below.

Commentary on the Four Purported Proofs

First, we make some general comments on interesting or unusual aspects of the four purported proofs. The theorem is true and one of the purported proofs is actually a proof of

[4] Because the purported proofs were constructed by undergraduate students and because the participants in this study were also undergraduate students, we will henceforth refer to the undergraduates in this study as "participants" to avoid confusion.

it, while three are not, that is, as proofs they are incorrect. In our previous paper, we provided a textual analysis -- a kind of line-by-line gloss or elaboration -- of the theorem and the four student-generated proofs that emphasized mathematical and logical points that a validator might, or might not, notice (Selden & Selden, 2003, pp. 10-18).

For example, we discussed such matters as the use of alternative terms (e.g., *assume* for *suppose*, *divides* instead of *multiple*), the role of individual sentences in furthering the argument, proper and improper uses of symbols, implicit assumptions (e.g., that a division of the argument into cases had been exhaustive), the correctness of inferences, computational errors, extraneous statements, and structural aspects of the four purported proofs. Global properties such as whether an argument proved the theorem, as opposed to some other theorem, were also noted.

Here we consider the four purported proofs, one-by-one, only very briefly. "Proof (a)" (Appendix 2, Phase 3) is not a proof. It consists of two independent subarguments each of which should have ended with the conclusion "n is a multiple of 3" or its equivalent, "n is divisible by 3." However, the odd case did not end this way, and the even case made this claim but did not properly justify it. In addition, while taking odd and even cases, when n^2 is a multiple of three, would not be wrong, these two cases seem a bit unusual. However, if each case were proved correctly, the proofs of the two cases would be essentially the same.

If "Proof (b)" is treated as a proof of the contrapositive, which one of the participants (CY) in the current study did, it would be peculiar to mention "the contrary" in the beginning. Under a contrapositive interpretation, the final step could have been omitted entirely or replaced by "Thus in either case n^2 is not a multiple of 3." We regard this as a proof of the theorem, although one that might have been written more clearly. Had the role of n^2 and the division into two independent subarguments (cases) been made explicit rather than implicit, the proof would have been less confusing for validators, especially inexperienced ones.

Given what the students who wrote the purported proofs knew, and what the participants in this study and the previous study knew, "Proof (c)" has a gap in the reasoning, although some mathematicians have pointed out that the result can be considered to be an immediate consequence of the definition of prime, so there is no gap. This observation by mathematicians points out the importance of context for validation. Accordingly, the fact that the participants in both studies had been told, at the beginning of the interviews, that the purported proofs were written by students like themselves, who were in a transition-to-proof course, gave them crucial contextual information. Indeed, several participants in the current study wondered what the students who wrote the purported proofs knew or had been allowed to assume.

"Proof (d)" begins with the conclusion and arrives at the hypothesis, although not in a straightforward way. Hence, it can be considered a proof of the converse of the stated theorem.

Results: Participants' Observed Validation Behaviors

Given that validation can be difficult to observe, it is remarkable how verbal and forthcoming the participants in this study were. This enabled us to gather a variety of data, much of which is presented and discussed below.

All participants appeared to take the task very seriously and some participants spent a great deal of time validating at least one of the purported proofs. For example, LH[5] initially took 25 minutes to validate "Proof (a)" before going on, and VL initially took 20 minutes to validate "Proof (b)". The minimum, maximum, and mean times for validating each

[5] Initials, like LH, designate individual participants.

purported proof are given in Table 1.

Table 1: Time (in minutes) taken initially to validate the purported proofs (during Phase 2)

	"Proof (a)"	"Proof (b)"	"Proof (c)"	"Proof (d)"
Maximum time	25	20	16	9
Minimum time	5	2	3	2
Mean time	8.8	8.5	6.3	4.5

The following validation behaviors[6] were observed as having been enacted by the participants; the percentages and absolute numbers are given in parentheses:

1. Underlined, or circled, parts of the purported proofs (100%, 16);
2. Pointed with their pencils or fingers to words or phrases, as they read along linearly (50%, 8);
3. Checked the algebra, for example, by "foiling" $(3n+1)^2$ (62.5%, 10);
4. Substituted numbers for n to check the purported equalities (37.5%, 6);
5. Reread all, or parts of, the purported proofs (87.5%, 14);
6. *Consulted the Fact Sheet to check something about multiples of 3 (56.25%, 9).*

Summarizing the above, participants used focus/reflection aids (1. & 2.); checked computations or tested examples (3. & 4.); revisited important points – perhaps as a protection against "mind wandering" (5.); and checked their own knowledge (6.). These actions all seem to be beneficial validation behaviors.

Results: Participants' Proffered Evaluative Comments

The participants sometimes voiced what they didn't like about the purported proofs. For example, CY objected to "Proof (b)" being referred to as a proof by contradiction. He insisted it was a contrapositive proof and twice crossed out the final words "we have a proof by contradiction". Fourteen (87.5%) mentioned the lack of a proof framework,[7] or an equivalent, even though they had been informed at the outset that the students who wrote the purported proofs had not been taught to construct proof frameworks.

Below are some additional features that seemed to bother some participants:

1. Lack of clarity in the way the purported proofs were written. Some referred to parts of the purported proofs as "confusing", "convoluted", "a mess", or not "making sense" (68.75%);
2. The notation, which one participant called "wacky";
3. The fact that "Proof (d)" started with n, then introduced m, and did not go back to n;
4. Not knowing what the students who had constructed the purported proofs knew or were allowed to assume:
5. Having too much, or too little, information in a purported proof. For example, one participant said there was "not enough evidence

[6] Ko and Knuth (2013, p. 27) referred to validation behaviors, such as checking line-by-line or example-based reasoning as "strategies" for validating proofs. We prefer the term "behaviors" as the act of underlining or circling parts of proofs is evidence of focus, not strategy, which usually entails a plan.

[7] A *proof framework* is a "representation of the 'top level' logical structure of a proof, which does not depend on a detailed knowledge of the mathematical concepts, but is rich enough to allow the reconstruction of the statement being proved or one equivalent to it." (Selden & Selden, 1995, p. 129). In practice, in this transition-to-proof course, this meant writing the hypotheses at the top of the nascent proof, leaving a blank space for the details, and writing the conclusion at the bottom of the proof, and perhaps, also unpacking the conclusion and writing as much as possible of the structure of the proof.

for a contradiction" in "Proof (b)";
6. The "gap" in "Proof (c)" which was remarked on by six participants.

Results: Individual Participants' Voiced Local and Overall Comments

Some participants made comments that indicated local concerns, but some comments were of an overall evaluative nature. That is, the overall comments often seemed to have more to do with making sense, having enough information, or being a "strong" proof, rather than with the structure of the purported proofs.

Indeed, no participant even commented on the strange division of "Proof (a)" into odd and even cases. This general lack of global, or overall, structural comments is similar to prior findings in the literature (e.g., Inglis & Alcock, 2012; Selden & Selden, 2003).

Local Comments

Some of the local comments on "Proof (a)" were:

MO: For the odd part, [I] don't like the string of equals.
[$n^2 = (3n + 1)^2 = 9n^2 + 6n + 1 = 3n(n + 2) + 1$]
KW: [It's] got a problem. $3n+1$, if $n=1$, is not odd – [it] would be even.
AF: This [$n^2 = 9n^2$] isn't equal.

Two of local comments on "Proof (b)" were:

FR: First off, [I am] not seeing the closing statement.
KK: [This is] not a proof because we don't introduce n, but we use n.

A sample local comment on the use of the universal quantifier in "Proof (c)" was:

CJ: [The bit about] where x is *any* integer worries me.

A local comment on the notation used in "Proof (d)" was:

LH: Why would you use m? ... [It's] kind of confusing with that m.

Overall Comments

Some overall comments on "Proof (a)" were:

CL: [It] needs more explanation -- I can't see where they are going.
CY: [The first case] doesn't seem right.
KW: [They are] not going where they need to go. …. No, not a proof. …
[I] don't think they've done what they need to do.
FR: I don't want to say this is done. Not a proper proof.
MO: [This is a] partial proof.

Three overall comments on "Proof (b)" were:

CL: Yes, [this one] looks a lot better. [It's] making more sense to me.
[than "Proof (a)"].

SS: [It's] not written well.
AF: [I] feel like it's a proof because [they're] showing that the two integers in between [i.e., $3n + 1$ and $3n + 2$] are not multiples of 3.

Four overall comments on Proof (c)" were:

CY: [I] just can't get my head around [it].
CJ: [I] need more information. [I] don't buy it.
KK: [This one is] closer [to a proof] than the others.
MO: I don't think (c) is a proof. [It] doesn't have enough information. [It] doesn't go into detail … [It] doesn't say exactly why it works.

Two sample overall comments on "Proof (d)" were:

MO: [He is] putting [in] more information than needs to be [there]. [This does] not help his proof.
LH: [This one's] not a strong proof.

Participants' comments, given in the above two sections, do not focus just on whether the theorem has been proved. They include evaluative comments about whether they liked the purported proofs, found them confusing or unclear in some way, or were lacking in some details or information. We suspect participants might have had difficulty separating matters of validity from matters of style and personal preference, or even from their own confusion, while reading the purported proofs.

Results: What the Participants Said They Do When Reading Proofs

In answer to the final debrief questions, all participants said that they check every step in a proof or read a proof line-by-line. All said they reread a proof several times or as many times as needed. All, but one, said that they expand proofs by making calculations or making subproofs. In addition, some volunteered that they work through proofs with an example, write on scratch paper, read aloud, or look for the framework. All of these actions can be beneficial.

In addition, ten (62.5%) said they tell if a proof is correct by whether it "makes sense" or they "understand it". These are cognitive feelings that, with experience, can be useful. Four (25%) said a proof is incorrect if it has a mistake, and four (25%) said a proof is correct "if they prove what they set out to prove." These last two views of proof call for some caution during implementation.

It is possible for a proof to have a minor mistake, perhaps a calculation error, that can be easily fixed, and hence, not "make sense" locally, but otherwise be correct. Indeed, it has been claimed that a past editor of the *Mathematical Reviews*[8] once said that "approximately one half of the proofs published in it were incomplete and/or contained errors, although the theorems they were purported to prove were essentially true." (de Villiers, 1990, p. 19). Consequently, it appears that, for most mathematicians, a mistake that is easily fixable does not mean the entire proof should be judged incorrect. Additionally, it is possible, especially for student proof attempts, to "end in the right place", but still have significant errors. Thus,

[8] *Mathematical Reviews* is a journal and online database published by the American Mathematical Society (AMS) that contains brief synopses, written and signed by mathematicians with appropriate expertise, of many published articles in mathematics, statistics, or theoretical computer science.

cognitive feelings such as those expressed in the previous paragraph need to be informed by appropriate proof construction and validation experiences.

Discussion and Teaching Implications

In answer to the initial research question, the participants in this study took their task very seriously, but made fewer final correct judgments (73% vs. 81%) than the undergraduates studied earlier (Selden & Selden, 2003) despite, as a group, being somewhat further along academically. In this study, 56% (9 of 16) of the participants were in their fourth year of university, whereas just 37.5% (3 of 8) of the undergraduates in our earlier study were in their fourth year. However, two of the eight participants in our earlier validation study had been able to prove the theorem themselves during Phase 1, whereas none of the 16 participants in this study proved the theorem.

Because the participants in this study were completing an inquiry-based transition-to-proof course emphasizing proof construction, in which validation had been modeled by the second author, we conjectured that they would be better at proof validation than those at the beginning of a transition-to-proof course (Selden & Selden, 2003), but they weren't. We have tentatively concluded that if one wants undergraduates to learn to validate "messy" student-constructed, purported proofs, in a reliable way, one needs to teach validation explicitly.

We stress this because it may seem counterintuitive. We note that, as students most mathematicians have received considerable implicit proof construction instruction through feedback on assessments and on their dissertations. However, most have received no explicit validation instruction, but are apparently very skilled at it.

There is at least one caveat regarding this study. It could be that we have been comparing "apples to oranges" as the participants in this study were not given a pre-test on proof validation, but instead are being compared to different students at the beginning of a transition-to-proof course at a different university. This could be remedied with another study that administered both a pre-test and a post-test consisting of validation items. However, it is difficult to imagine a course that gave more attention to helping students with proof construction or that demonstrated validation of proofs more explicitly than this one, with such limited effect on students' validation behaviors.

As to how one might possibly teach students to validate "messy" student-constructed proofs, Boyle and Byrne (2014, Table 1) have suggested a rubric, which they refer to as a "proof assessment tool" meant to help university teachers give formative feedback to students on their proofs. Byrne is currently using this rubric to have her transition-to-proof course students comment on each other's proof attempts (personal communication, March 8, 2014). It will be interesting to see whether Byrne needs to give her students explicit instruction on how to use this rubric or whether her students can use it "off the shelf" without explicit instruction, and also how proficient her students become at giving helpful feedback to each other on proofs and on validation.

We conjecture, as a result of our current study, that it's not enough for professors to model validation, as was done by the second author, and for students to observe those validations, as our participants did, rather students also need to practice validation, which is a reason to hope that Byrne will be successful with her current students.

Future Research

In addition to proof validation, there are three additional related concepts in the literature: proof comprehension, proof construction, and proof evaluation. There has been little research on how these four concepts are related. In this study, we investigated one of these relationships -- whether improving undergraduates' proof construction abilities would enhance their proof validation abilities and have obtained some negative evidence.

Proof comprehension means understanding a (textbook or lecture) proof. Mejia-Ramos, Fuller, Weber, Rhoads, and Samkoff (2012) have given an assessment model for proof comprehension, and thereby described proof comprehension in practical terms. Examples of their assessment items include: Write the given statement in your own words. Identify the type of proof framework. Make explicit an implicit warrant in the proof. Provide a summary of the proof.

Proof construction means constructing correct proofs at the level expected of mathematics students (depending which year they are in their program of study).

Proof evaluation was described by Pfeiffer (2011) as "determining whether a proof is correct and establishes the truth of a statement (validation) and also how good it is regarding a wider range of features such as clarity, context, sufficiency without excess, insight, convincingness or enhancement of understanding." (p. 5).

While it is still an open question as to how these four concepts are related, in addition to our study, Pfeiffer (2011) conjectured that practice in proof evaluation could help undergraduates appreciate the role of proofs and also help them in constructing proofs for themselves. She obtained some positive evidence, but her conjecture needs further investigation. As for proof comprehension, it is an open question as to whether practice in proof comprehension would help any of proof evaluation, proof validation, or proof construction.

In addition, there is anecdotal evidence, obtained from several mathematics department chairpersons, that some of today's transition-to-proof courses/textbooks are thought to be inadequate for the task of actually transitioning students from lower-level undergraduate mathematics courses to upper-level undergraduate proof-based mathematics courses, such as abstract algebra and real analysis. Whether this is the case, and to what degree, should be investigated.

Finally, we feel that there is a need to develop characteristics of a reasonable learning progression for tertiary proof construction, going from *novice*[9] (lower-division mathematics students) to *competent* (upper-division mathematics students), on to *proficient* (mathematics graduate students), and eventually to *expert* (mathematicians).

References

Bills, L., & Tall, D. (1998). Operable definitions in advanced mathematics: The case of The least upper bound. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 2 (pp. 104-111). Stellenbosch, South Africa: University of Stellenbosch.

Boyle, J., & Byrne, M. (2014). Supporting students to construct proofs: An argument assessment tool. *Proceedings of the 17th Annual Conference on Research in Undergraduate Mathematics Education*.

Coppin, C. A., Mahavier, W. T., May, E. L., & Parker, G. E. (2009). *The Moore Method: A pathway to learner-centered instruction* (MAA Notes No. 75). Washington, DC: MAA.

de Villiers, M. (November, 1990). The role and function of proof in mathematics. *Pythagoras*, *24*, 17-24.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the age of the computer*. Oxford: Blackwell.

[9] The terms *novice*, *proficient*, *competent*, and *expert* have been adapted from the Dreyfus and Dreyfus (1986) novice-to-expert scale of skill acquisition.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*(4), 358-390.

Ko, Y.-Y., & Knuth, E. J. (2013). Validating proofs and counterexamples across content domains: Practice of importance for mathematics majors. *Journal of Mathematical Behavior*, *32*, 20-35.

Mahavier, W. S. (1999). What is the Moore Method? *PRIMUS*: Problems, *Resources*, *and Issues in Mathematics Undergraduate Studies*, *9*, 339-354.

Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*(1), 3-18.

McKee, K., Savic, M., Selden, J., & Selden, A. (2010). Making actions in the proving process explicit, visible, and 'reflectable.' In *Proceedings of the 13th Annual Conference on Research in Undergraduate Mathematics Education*. Available at http://sigma.maa.org/rume/crume2010/Archive/McKee.pdf.

Pfeiffer, K.(2011). Features and purposes of mathematical proofs in the view of novice students: Observations from proof validation and evaluation performances. Doctoral dissertation, National University of Ireland, Galway.

Selden, A., McKee, K., & Selden, J. (2010). Affect, behavioural schemas and the proving process. *International Journal of Mathematical Education in Science and Technology*, *41*(2), 199-215.

Selden, A., & Selden, J. (2013). The genre of proof. In M. N. Fried & T. Dreyfus (Eds.), *Mathematics and Mathematics Education: Searching for Common Ground*. New York: Springer, Advances in Mathematics Education series, 2013. Also, available as Tennessee Technological University Mathematics Department Technical Report No. 2013-1, which can be downloaded from http://www.tntech.edu/math/reports/.

Selden, A., & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem*? Journal for Research in Mathematics Education*, *34*(1), 4-36.

Selden, J., & Selden, A. (2009). Teaching proving by coordinating aspects of proofs with students' abilities. In D.A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and Learning Proof Across Grades: A K-16 Perspective* (pp. 339-354). New York/Washington, DC: Routledge/National Council of Teachers of Mathematics.

Selden, J., & Selden, A. (2008). Consciousness in enacting procedural knowledge. *Proceedings of the 11th Annual Conference on Research on Mathematics Undergraduate Mathematics Education.* Retrieved from http://rume.org/crume2008/Proceedings/Selden_Procedural%20LONG.pdf.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*, 123-151.

Weber, K., & Alcock, L. (2005). Warranted implications to understand and validate proofs. *For the Learning of Mathematics*, *25*(1), 34-38 & 51.

Weber, K. (2008). How do mathematicians determine if an argument is correct? *Journal for Research in Mathematics Education*, *39*(4), 431-439.

Appendix 1: Fact Sheet

[from Selden and Selden (2003), p. 32]

FACT 1. The positive integers, $\mathbf{Z}^+$, can be divided up into three kinds of integers -- those of the form $3n$ for some integer n, those of the form $3n + 1$ for some integer n, and those of the form $3n + 2$ for some integer n.

For example,

1,	2,	3,	4,	5,	6,	7,	8,	9,	10,	11…
		$3n$	$3n+1$	$3n+2$	$3n$	$3n+1$	$3n+2$			
			where $n = 1$			where $n = 2$				

FACT 2. Integers of the form $3n$ (that is, 3, 6, 9, 12, . . .) are called *multiples of* 3.

FACT 3. No integer can be of two of these kinds simultaneously. So m is not a multiple of 3 means the same as m is of the form $3n+1$ or $3n+2$.

Appendix 2: Interview Protocol

[from Selden and Selden (2003), pp. 32-33]

PHASE 1: 'Warm Up' Exercises

For any positive integer n, if n^2 is a multiple of 3, *then n is a multiple of* 3.

1. Explain, in your own words, what the above statement says.
2. Give some examples of the above statement.
3. Does the above statement seem to be true? How do you tell?
4. Do you think you could give a proof of the above statement?

PHASE 2: Sequential consideration of 'Proofs' (a), (b), (c), (d). [The purported proofs were presented to the participants, one page at a time, during this Phase. The purported proofs are given below under Phase 3.]

PHASE 3: 'Recap' on the 'Proofs'

Below are several purported proofs of the following statement:

For any positive integer n, if n^2 is a multiple of 3, *then n is a multiple of* 3.

For each one, decide whether or not it is a proof. Try to "think out loud" so you can let me in on your decision process. If it is not a proof, point out which part(s) are problematic. If you can, say where, or in what ways, the purported proof has gone wrong.

(a). Proof: Assume that n^2 is an odd positive integer that is divisible by 3. That is $n^2 = (3n + 1)^2 = 9n^2 + 6n + 1 = 3n(n + 2) + 1$. Therefore, n^2 is divisible by 3. Assume that n^2 is even and a multiple of 3. That is $n^2 = (3n)^2 = 9n^2 = 3n(3n)$. Therefore, n^2 is a multiple of 3. If we factor $n^2 = 9n^2$, we get $3n(3n)$; which means that n is a multiple of 3. ■

(b). Proof: Suppose to the contrary that n is not a multiple of 3. We will let $3k$ be a positive integer that is a multiple of 3, so that $3k + 1$ and $3k + 2$ are integers that are not multiples of 3. Now $n^2 = (3k + 1)^2 = 9k^2 + 6k + 1 = 3(3k^2 + 2k) + 1$. Since $3(3k^2 + 2k)$ is a multiple of 3, $3(3k^2 + 2k) + 1$ is not. Now we will do the other possibility, $3k + 2$. So, $n^2 = (3k + 2)^2 = 9k^2 + 12k + 4 = 3(3k^2 + 4k + 1) + 1$ is not a multiple of 3. Because n^2 is not a multiple of 3, we have a contradiction. ■

(c). Proof: Let n be an integer such that $n^2 = 3x$ where x is any integer. Then $3|n^2$. Since $n^2 = 3x$, $nn = 3x$. Thus $3|n$. Therefore if n^2 is a multiple of 3, then n is a multiple of 3. ■

(d). Proof: Let n be a positive integer such that n^2 is a multiple of 3. Then $n = 3m$ where $m \in \mathbf{Z}^+$. So $n^2 = (3m)^2 = 9m^2 = 3(3m^2)$. This breaks down into $3m$ times $3m$ which shows that m is a multiple of 3. ■

PHASE 4. Final Questions

1. When you read a proof is there anything different you do, say, than in reading a newspaper?
2. Specifically, what do you do when you read a proof?
3. Do you check every step?
4. Do you read it more than once? How many times?
5. Do you make small subproofs or expand steps?
6. How do you tell when a proof is correct or incorrect?
7. How do you know a proof proves this theorem instead of some other theorem?
8. Why do we have proofs?

AN ANALYSIS OF TRANSITION-TO-PROOF COURSE STUDENTS' PROOF CONSTRUCTIONS WITH A VIEW TOWARDS COURSE REDESIGN

John Selden — New Mexico State
Ahmed Benkhalti — New Mexico State
Annie Selden — New Mexico State

The purpose of this study was to gain knowledge about undergraduate transition-to-proof course students' proving difficulties. We analyzed the final examination papers of students in one such course. Our perspective included drawing inferences about students' sometimes automated links between situations and mental, as well as physical, actions. We have identified process, rather than mathematical content, difficulties such as not constructing a proof framework, not unpacking the conclusion, and not using definitions correctly. The ultimate goal is to contribute to an understanding of some of these kinds of difficulties as pedagogical content knowledge with which to teach or redesign transition-to-proof courses.

Key words: Transition-to-proof, Proof construction, Pedagogical content knowledge, Actions, Proof framework

This paper presents an analysis of transition-to-proof course students' final examinations in an effort to describe some of their main proving difficulties. By inferring kinds of difficulties in students' proof construction processes from their written proof attempts, and by focusing away from specific fields of mathematics, we begin to answer the question: How can the general proving process be taught so that it applies broadly to many fields of mathematics? For example, what knowledge, habits of mind, and self-efficacy will facilitate students' proof construction processes and might be candidates for explicit teaching in a transition-to-proof course?

Analyzing student examination papers in the service of teaching and course redesign might seem unpromising for a typical, that is, content driven, mathematics course because there is already mathematical terminology to connect student examination difficulties with parts, or precursors, in a way useful in teaching. For example, consider student difficulties in finding $d/dx \sin x^2$. The chain rule and composition of functions come immediately to mind as things the student must be able to use.

In contrast, the proving difficulties we examined were mostly about the *process* of constructing a proof, not mathematical content. Thus new concepts and vocabulary may emerge to connect overall difficulties, for example, not being able to finish constructing a particular complete proof, with contributing parts or precursors, for example, the ability to use abstract definitions, in a way useful for teaching and course design. We turn now to some related prior research.

Literature Review

While some studies of students' proving difficulties have been conducted before, they have not been so closely aimed at course design, especially design based on the process of proof construction. Also, several studies have been conducted with students who were mathematically more advanced than ours. For example, Selden and Selden (1987) examined errors and misconceptions in undergraduate abstract algebra students' proof attempts. However, the difficulties reported there have little in common with those observed in this study. In addition, Weber's (2001) study, contrasting undergraduate abstract algebra students with doctoral students in algebra, showed that the latter had strategic (content) knowledge to use in constructing abstract algebra proofs that the undergraduates did not have. Our study, in contrast, gives insight into the proving difficulties of relative beginners, that is, undergraduate students at the end of a transition-to-proof course. We note that Moore (1994) observed a traditionally taught transition-to-proof course and reported seven student proving difficulties, some of which do overlap with our categories, although in general, our categories are more fine-grained. In addition, Baker and Campbell (2004) reported three observations of somewhat less sophisticated transition-to-proof course students. Selden and Selden (1995) did observe process difficulties in unpacking the logic of informal mathematical statements. They reported that informal statements, that is, those that departed from the simplest natural language rendering of predicate and propositional calculus were difficult for students to unpack and hence difficult to prove. This information was indeed used in designing our current course. In it we decided to write mathematical statements in the course notes rather formally so the students would not need to unpack their logical structure and could focus on the rest of the proof construction. This is not because we did not value learning to autonomously unpack the logical structure of statements, but because certain proving actions seem to call for self-efficacy (Selden & Selden, 2014), which can be encouraged by providing students early opportunities to succeed in constructing proofs. Thus, we gave building self-efficacy priority (Selden & Selden, 2014).

There is some additional literature that supports our theoretical perspective, but it is perhaps best understood in that section below.

Theoretical Perspective

First, we will suggest some psychologically-based ideas and then mention a few concepts that have emerged from earlier iterations of the course under consideration. We view the process of proof construction as a *sequence of mental* (e.g., "unpacking" the meaning of the conclusion in inner speech) *or physical* (e.g., drawing a diagram) *actions*. Such a sequence of actions is somewhat related to, and extends, what we have earlier called a "possible construction path" of a proof, illustrated in Selden and Selden (2009). The actions derive from a person's nonobservable, and sometimes partly nonconscious, *inner interpretation* of usually outer and observable situations in a partly completed proof construction.

Inner interpretations cannot be observed, but they can be inferred, sometimes very convincingly. Norton and D'Ambrosio (2008, pp. 14-15) provide an illustration of this for two middle school students, Will and Hillary, who viewed the same external situation involving a fraction such as 2/3. Hillary had (in her knowledge base) a partitive fractional scheme, as well as a part whole fractional scheme, while Will had only the second scheme. This caused Will and Hillary to "see" the external situation differently, that is, to have differing inner interpretations, and hence to act differently. Hillary was able to solve the problem, but Will couldn't. Will could only solve the problem after he had developed a partitive fractional scheme, and presumably then experienced a richer inner interpretation.

When several similar (inner) situations are followed by several similar actions, an *automated link* may be learned between such situations and actions. A situation is then followed by an action, without the need for any conscious processing between the two (Selden, McKee, and

Selden, 2010). This appears to be a form of procedural memory, which can be thought of as "knowing how" as opposed to "knowing that". A brief discussion of kinds of memory, including procedural, can be found in Ranganath, Libby, and Wong (2012, pp. 121-123). Also, automated actions have been extensively studies by psychologists interested in their occurrence in everyday life (Bargh, 2014, 1997).

Many proving actions appear to be the result of the enactment of small, linked, automated situation-action pairs that we have termed *behavioral schemas* (Selden, McKee, & Selden, 2010; Selden & Selden, 2008). Automating actions can considerably reduce the burden on working memory, a very limited resource, and thus tends to reduce errors. (Baddeley, 2000) The value of automating actions in proof construction is illustrated in the comments following Sample Corect Proof 3, below.

Nonemotional cognitive feelings and *nonconscious priming* can influence whether a *situation-action* link is activated. Nonemotional cognitive feelings, such as the feeling of being, or not being, on the right track, typically are vague conscious states that pervade one's whole conscious field and can combine with anything being focused upon (Selden, McKee, & Selden, 2010). Nonconscious priming occurs when an individual is unaware of the way a situation is influencing an action. For example, a student attempting to construct a proof might have written several line into a proof which, although true, do not move the proof forward. The student might then wrongly decide the proof was finished without realizing those several lines made the work "look like" something useful had been done. Below we discuss this appearing to occur in one student's work. (See Sample Incorrect Student Proof Attempt 2.)

Some actions can be *meta-actions*, that is, actions on one's own cognition, such as focusing on a particular part of a partly finished proof. Meta-actions should be distinguished from meta-cognition, that is, thinking about one's own thinking. Retrospective meta-cognition is likely to be a useful addition to understanding one's own proof construction, but attempting simultaneous meta-cognition could compete for working memory with the cognition that it is observing.

Some actions are beneficial for proof construction and should be initiated or encouraged. (See Sample Incorrect Student Proof Attempt 4 below for a beneficial action, not taken, that should be encouraged.) Other actions can be detrimental and should be eliminated or discouraged.

During proof construction, the partly completed proof and scratchwork are an important part of the construction. They can be used as aids to reflection and to reduce the burden on working memory. For a few psychologists, these might even be seen as an external part of cognition, which is normally seen as entirely mental and inner. In this setting, a proof should be regarded as a *text*, that is, as something that can be passed between persons. It consists of *some* of the actions in a proof construction. Indeed, this suggests why it is often hard for a student to mimic the proof of a previous theorem in trying to prove another theorem. If anything could be mimicked, it would be some of the actions taken during the proof construction of the previous theorem, many of which are not available for later viewing.

When students are first learning proof construction, many actions, such as the construction of a proof framework (Selden & Selden, 1995), can be automated. A good way to learn such actions is probably through "coached experience", like riding a bicycle or playing soccer. This is why we regularly engage students in proof co-construction during some of our transition-to-proof classes.

We now turn away from the more psychological part of our perspective to some more mathematical aspects that have emerged from teaching several earlier iterations of the current course.

We first describe the writing of a proof framework in more detail than can be found in Selden and Selden (1995). Proof frameworks are determined by the logical structure of what is to be proved. The most common form of theorem in our course notes is: some quantified variables; then "if P", where P is a predicate about those variables; then "then Q", where Q is another predicate involving some of the variables.

A *proof framework* starts by introducing the variables. If "for all $a \in A$" occurs in the theorem, one writes in the emerging proof "Let $a \in A$", in which case a is henceforth regarded as fixed, but unspecified. If "there is a $b \in B$" occurs, one must "find/create" such a b and a space is left to insert or explain that. If the quantifiers are mixed, some "for all" and some "there exist", then in the proof these should be introduced in the same order as in the statement of the theorem. This avoids inadvertently changing the meaning of the theorem during the proving process. Where the theorem says "if P", one writes in the proof, "Suppose P" and leaves a space for further parts of the proof. Where the theorem says "then Q", one writes at the end of the emerging proof "Therefore Q". This produces the *first-level* of the proof framework.

At this point the student should focus on Q and "unpack" its meaning, that is, remember or look up its definition, being careful to change the names of its variables to fit the proof at hand. It may happen that the meaning of Q has the same logical form as the original theorem. In that case, one can repeat the above process, providing a *second-level* of proof framework which is written into the blank space immediately above "Therefore Q". If in writing the second-level framework, some variables have already been introduced, one does not re-introduce them.

All of this is rather complicated to explain, but much easier to understand in practice, and is illustrated in our sample proofs, below. Also, once students can produce and use a proof framework for the above "if P, then Q" logical structure, it appears to be relatively easy to introduce frameworks for the seven or so other logical structures needed in the course. Finally, we are not claiming that mathematicians write proofs in the way we are describing, but only that doing so will be helpful for students and that mathematicians will accept the results.

We turn now to the idea of *operable interpretation*s of definitions. Consider the following definition: Let $f: X \rightarrow Y$ be a function and $A \subseteq X$. Then $f(A) = \{ y \mid y \in Y$ and there is $a \in A$ so that $f(a) = y\}$. To use this, if one knows "$q \in f(A)$" one can say "there is $p \in A$ so that $f(p) = q$". Also, if one knows "$p \in A$" then one can say "$f(p) \in f(A)$". One might expect that a beginning transition-to-proof course student would be able to autonomously discover such operable interpretations of definitions, but we have noticed that many cannot. Learning to do so would be a useful skill for anyone wishing to read mathematics or prove theorems independently. However, before this skill is learned, it may be helpful to provide some of the operable interpretations. There are around 30 in our course.

Searching the course notes or one's own knowledge base may not seem to be a very sophisticated skill, but we find that some students do not do it when some result or definition is called for in constructing a proof. We now try to arrange the course notes so that students can experience the benefits of noticing useful prior results.

We mean by *exploration* in proof construction doing something of unknown value, for example, finding or constructing new "objects" or "manipulating" them. A reltively easy example can be seen below in the middle of Sample Correct Proof 1, where the prover cannot know that manipulating $abab$ or $abab = e$ will be helpful. We suggest that exploration is aided by a students' *self-efficacy* (Selden & Selden, 2014) and for this reason we try to arrange for students to have early proving successes in our course.

We will now borrow a point from the *genre of proof*, namely, that definitions available outside of a proof are not normally written into it, at least not in proofs published in journals (Selden & Selden, 2013). Quoting an entire definition exactly into a proof can wrongly suggest to a student that something has been done that moves the proof forward. The student may then prematurely stop work on the proof. This should not be confused with *using* a definition in a proof. For example, in using the fact that f is continuous in a proof, one normally writes, "Because f is continuous, there is a δ such that …". This looks rather like, but is not, quoting the definition.

Finally, we have found it helpful to have a, at least crude, gauge of the difficulty of a proof, independent of the ideas in the rest of this perspective. We say proofs are of *Type 1, 2, or 3* as follows. A *Type 1 proof* calls for a student to see the need for a lemma, a subproof that could be proved separately, but could also be found located in the course notes. In a *Type 2 proof*, the student must articulate a lemma not proved in the notes, but the lemma's articulation and its proof are straightforward. In a *Type 3 proof*, either the articulation or proof is not straightforward and may require insight or exploration. (Selden & Selden, 2013b, pp. 319-320).

The Course

The course, from which the data came, was *inquiry-based* as regards the proofs, but not as regards the mathematical structures or theorems. It was taught entirely from notes with students constructing original proofs and receiving critiques in class. The one-semester three-credit course is meant as a second-year university transition-to-proof course for mathematics and secondary education mathematics majors. It was given at a Southwestern Ph.D.-granting university and was taught in a very modified Moore Method way (Coppin, Mahavier, May, & Parker, 2009; Mahavier, 1999). That is, students were given course notes with definitions, questions, requests for examples, and statements of theorems to prove.

The students in this study proved the theorems outside of class and presented their proofs in class on the blackboard and received extensive critiques. These critiques consisted of careful line-by-line readings and validations of the students' proof attempts. This was followed by a second reading of the students' proof attempts, indicating how these might have been written in "better style" to conform to the genre of proofs (Selden & Selden, 2013a). Once these corrections and suggestions had been made, the student, who had made the proof attempt, was asked to write it up carefully, including any corrections and suggestions, for duplication for the entire class. In this way, by the end of the semester, the students had obtained one correct, well-written proof for each theorem in the course notes. Sometimes, if the students seemed to need it, there were mini-lectures on topics such as logic or proof by contradiction.

The homework, assigned each class period, consisted of requests for proofs of the next two or three theorems in the course notes. These proof attempts were handed in at the beginning of the next class to the third author, who determined "on the spot", based on the students' written work, which students would be asked to present their proof attempts on the blackboard that day. In addition to presenting their attempted proofs in class, the students had both take-home and in-class final examinations, each of which consisted of four theorems, new to them, to prove. The mathematical topics considered in the course included sets, functions, continuity, and beginning abstract algebra in the form of a few theorems about semigroups and homomorphisms. However, the teaching aim was to facilitate students' learning the proof construction *process* in an embodied way through experience constructing as many different kinds of proofs as possible, especially in abstract algebra and real analysis, and *not* to learn a particular mathematical content.

Methodology

Guided by our theoretical perspective and our aim for the course, we analyzed all 16 four-proof take-home, and all 16 four-proof in-class, final examination papers from the course. Altogether 128 student proof attempts were analyzed in detail through several iterations, using a combination of grounded theory and textual analysis. When we found indications of a difficulty in a student's attempted proof, we drew inferences about the probable proving action that might have led to that difficulty. We were looking for categories at a level of abstraction above specific mathematical topics so they would reflect *process* difficulties. For example, we considered a student's not unpacking a conclusion, as opposed to a student having difficulty with a particular mathematical concept, such as a minimal ideal in a semigroup. We began with no particular categories in mind and made several passes through the data, until we came to an agreement on what we saw in each student's proof attempt. In this way our categories emerged from our data.

We expected, and found, that when one difficulty occurred in a student's proof attempt, other difficulties often also occurred. In addition, since our main interest was in finding a few difficulties that we might be able to alleviate with explicit teaching interventions, we did not attempt to search for categories that did not overlap or were not within other categories. Such information might be useful in designing teaching interventions.

Categories: The Most Common Student Difficulties

We have thus far identified the following categories: omitting beneficial actions; taking detrimental actions; inadequate proof framework (e.g., not unpacking the conclusion); mathematical syntax errors; wrong or improperly used definitions; misuse of logic; insufficient warrant; assumption of all or part of the conclusion; extraneous statements; assumption of the negation of a previously established fact; difficulties with proof by contradiction; inappropriately mimicking a prior proof; mathematical syntax errors, failure to use cases when appropriate; incorrect deduction; assertion of an untrue result; and computational errors.

While most categories can be easily understood from their names, there is one sufficiently odd that it might benefit from an illustration. Here is an example of a mathematical syntax error. In an attempt to prove that the split domain function h, defined by $h(x) = f(x)$ if $x \geq a$ and $h(x) = g(x)$ if $x < a$, is continuous at a, given that both f and g are continuous at a and $f(a) = g(a)$, one student wrote: "$|f(x)-f(a)|< \varepsilon/2 - |g(x)-g(a)|< \varepsilon/2$". This action, subtracting a statement such as "$|g(x)-g(a)|< \varepsilon/2$", from another statement, violates normal mathematical syntax. Subtraction is an arithmetic operation used between numbers or variables representing numbers, not a logical operation used between statements.

In our textual analysis below, we illustrate omitting beneficial actions; taking detrimental actions; inadequate proof frameworks; not unpacking the conclusion; and extraneous statements (e.g., writing a definition that can be found outside of a proof into it).

Textual Analysis of Sample Correct Proofs and Corresponding Sample Incorrect Student Proof Attempts

In the following section, we consider both a sample correct, and a corresponding sample incorrect student proof attempt, of the same four theorems. We are numbering the lines with bold square brackets for the purpose of referencing them when we comment on them.

Sample Correct Proof 1. The first theorem we consider is: *Theorem. Let S be a semigroup with an identity element e. If, for all s in S, ss = e, then S is commutative.* Our sample correct proof is given below.

Proof:
[1] Let S be a semigroup with identity e.
[2] Suppose for all $s \in S$, $ss = e$.
[3] Let a, b be elements in S.
[4] Now $abab = e$, so $(abab)b = eb = b$.
[5] But $(abab)b = aba(bb) = (aba)e = aba$.
[6] Thus $aba = b$, so, $(aba)a = ba$, and $(aba)a = ab(aa) = abe = ab$.
[7] Thus $ba = ab$.
[8] Therefore, S is commutative. QED.

We imagine that an idealized student prover would first write the hypotheses **[1]** and **[2]**, leave a space for the body of the proof, and then write the conclusion **[8],** thereby completing the first-level proof framework. Next our idealized prover would unpack the conclusion **[8]**, perhaps using scratchwork, and if necessary, consult the definition of commutative, which is in the course notes. By doing so, our idealized prover would know that he/she has to introduce two arbitrary elements of the semigroup, say a and b **[3]**. Then the prover could write line **[7]**, thereby completing the second-level proof framework. What is required next is some "exploring", that is, some manipulations, that prover cannot know will be useful, until lines **[4]**, **[5]**, and **[6]** can be written.

Sample Incorrect Student Proof Attempt 1. Everything is reproduced below as written by the student, including the student's scratchwork, except for the line numbers.

Proof:
[1] Let S be a semigroup with an identity
element, e. **[2]** Let $s \in S$ such that $ss = e$.
[3] Because e is an identity element, $es = se = s$.
[4] Now, $s = se = s(ss)$.
[5] Since S is a semigroup, $(ss)s = es = s$.
[6] Thus $es = se$.
[7] Therefore, S is commutative. QED.

SCRATCHWORK:
7.1: A semigroup is called commutative or Abelian if, for each a and $b \in S$, $ab = ba$.
7.5: An element e of a semigroup S is called an identity element of S if, for all $s \in S$, $es =$
$se = s$.

We assume the student wrote the first-level proof framework at the start, lines **[1], [2]** and **[7]**. Line **[2]**, as written, only hpothesizes a single element s so that $ss = e$. Line **[2]** should have been "Suppose for all $s \in S$, $ss = e$." With this change, the first-level framework would have been correct. Also, we cannot be sure line **[7]** was written before the rest of the proof. If the student did not write all of lines **[1], [2]** and **[7]** first, this would constitute a beneficial action not taken.

In addition, despite being aware of the definition of Abelian written in the scratchwork, the student did not write the second-level framework by introducing arbitrary a and b at the top, followed by "Then $ab = ba$" right above the conclusion. Had the student written the correct second sentence in line **[2]** and taken these two actions, the situation would have been appropriate for exploring and manipulating an object such as $abab$. We think that such exploration calls for some self-efficacy, but can lead to a correct proof.

Line **[3]** violates the mathematical norm of not including in the proof definitions that can easily be found outside the proof. Also, this does not move the proof forward. The next three lines **[4]**, **[5]**, and **[6]** are not wrong, but also do not move the proof forward because to prove commutativity, one needs two arbitrary elements. These actions not only do not move the proof forward, but might have been detrimental. Through non-conscious priming, they might have wrongly convinced this student that he/she had accomplished something and prematurely brought work on the proof to an end.

Sample Correct Proof 2. Next we consider the following: *Theorem. Let S and T be semigroups and f:S→T be a homomorphism. If G is a subset of S and G is a group with identity e, then f(G) is a group.* Our sample correct proof is given below.

Proof:
[1] Let S and T be semigroups and $f{:}S \to T$ be a homomorphism.
[2] Let G be a subset of S and G be a group with identity e.

Part 1. [3] Note that G is a subsemigroup of S so, by Theorem 20.4, $f(G)$ is a semigroup.

Part 2. [4] Let $y \in f(G)$. **[5]** Then there is $x \in G$ so that $f(x) = y$. **[6]** Now $f(e) \in f(G)$ and $f(e)y = f(e)\,f(x) = f(ex) = f(x) = y$. **[7]** Similarly, $y\,f(e) = y$. **[8]** Thus, $f(e)$ is an identity for $f(G)$.

Part 3. [9] Let q in $f(G)$. **[10]** Then there is $p \in G$ so that $f(p) = q$. **[11]** Now because G is a group, there is $p' \in G$ so that $pp' = p'p = e$. **[12]** Thus $q\,f(p') = f(p)\,f(p') = f(pp') = f(e)$, and **[13]** $f(p')q = f(p')\,f(p) = f(p'p) = f(e)$. **[14]** Thus, each $q \in f(G)$ has an inverse, $f(p')$, in $f(G)$.

[15] Therefore, is a $f(G)$ group. QED.

An idealized student prover would first write lines **[1]** and **[2]**, then leave a space for the body of the proof, and finally write the conclusion **[15]** at the bottom. He/she would next unpack the conclusion, and if necessary, look up the definition of group. This would lead our prover to consider three parts, namely, **[3]** $f(G)$ is a subsemigroup of T, **[4]** $f(e)$ is an identity for $f(G)$, and **[14]** each $q \in f(G)$ has an inverse in $f(G)$.

Next our idealized prover can begin on Part 1. But this is almost immediate because of a previous theorem in the course notes that says that the homomorphic image of a semigroup is a semigroup. Hence, line **[3]**.

Our idealized prover can then go on to prove Part 2. For this, he/she would need to use the meaning of the definition of an identity and consider an arbitrary element of $f(G)$ **[4]**, and have to conjecture that the identity of $f(G)$ is $f(e)$, the image of the identity e of G. This would lead to using the meaning of $x \in$ f(G) and line **[5].** Then using the meaning of homomorphism would give line **[6]**, showing $f(e)$ is a left identity for G. Then line **[7]** would follow by similarity and line **[8]** would conclude a proof of Part 2.

Lastly, our prover would work on Part 3. For this, he/she would use the meaning of inverse element and consider an arbitrary element q in $f(G)$. He/she would then call on the meaning of $q \in f(G)$ to notice **[10]** that q can be written as $f(p)$ for some p in G and, using the meaning of homomorphism, show that the image of the inverse of p is the inverse of q by lines **[11]**, **[12]**, and **[13].** Line **[14]** asserts the conclusion of Part 3, and the proof is complete, according to the meaning of the definition of group.

We observe that the above idealized student prover often unpacked the meaning of a definition, by using what we have called its operable interpretation, and then altering the names of the variables to fit the theorem at hand. We also note that actions were rarely warranted, as is customary, as if they were completely transparent. However, for some beginning students, perhaps even many, we have found that using definitions in this way is not at all transparent.

Sample Incorrect Student Proof Attempt 2. We next consider a sample student proof attempt of the same theorem.

Proof:
[1] Let S and T be semigroups and $f{:}S{\to}T$ be a homomorphism.
[2] Suppose $G \subseteq S$ and G is a group with identity e.
[3] Since G is a group and it has identity e, then for each element g in G
there is an element g' in G such that $gg' = g'g = e$.
[4] Since f is a homomorphism, then for each element $x \in S$ and $y \in S$, $f(xy)=f(x)f(y)$.
[5] Since $G \subseteq S$, then $f(gg')=f(g)f(g')$. So $f(gg') = f(g'g) = f(e)$.
[6] So $f(G)$ has an element $f(e)$ since f is a function.
[7] Therefore, $f(G)$ is a group. QED.

The student has written the first-level framework correctly, lines **[1]**, **[2]** and **[7]**, that is, assuming the last line **[7]** was written immediately after writing the first two lines. To complete the framework, the student should have next considered $f(G)$ and noted that there are three parts to prove, as indicated in the sample proof above. These are beneficial actions the student did not take.

Instead, the student wrote into the proof the definition of G being a group **[3]** and f being a homomorphism **[4]**. These are actions that do not move the proof forward and are detrimental because they can convince the student that something useful has been done. Perhaps the student was trying to show the existence of an identity and inverses for $f(G)$ in lines **[5]** and **[6]** and was unsuccessful, but we cannot know this.

This student's work may suggest that he/she had some intuitive grasp of the concepts involved, and it may be tempting to give partial credit to the student. But from the point of view of having a student learn to *construct* proofs, doing so may send the "wrong message".

Further, this student's work is reminiscent of Carrisa, who was attempting prove or disprovethe statement: *Let* ϕ *be a* 1-1 *homomorphism from* (G,o) to $(H,*)$. If *G is an abelian group, then H is an abelian group*. Carrisa concentrated on elements of G, the wrong place to start, and mistakenly said the statement was true. She ignored, or did not see, the fact that ϕ is not known to be onto (Melhuish, 2014, pp. 3-4). Had she written a complete framework for a proof of H being a group or H being Abelian, she might have seen that she had inadequate information to finish a proof.

Sample Correct Proof 3. Next we consider the following: *Theorem. If A, B, and C, are sets and* $C\backslash B \subseteq C\backslash A$, *then* $C \cap A \subseteq C \cap B$. A sample correct proof follows.

Proof:
[1] Let A, B, and C be sets. **[2]** Suppose $C\backslash B \subseteq C\backslash A$. **[3]** Let $x \in C \cap A$. **[4]** Suppose $x \notin B$.
[5] Then $x \in C\backslash B$, **[6]** so $x \in C\backslash A$. **[7]** Thus $x \notin A$. **[8]** This is a contradiction.
[9] So $x \in B$ and **[10]** thus $x \in C \cap B$. **[11]** Therefore, $C \cap A \subseteq C \cap B$. QED.

An idealized student prover would first construct the first-level proof framework **[1]**, **[2]** and **[11]**, then "unpack" the conclusion, that is, use the operable interpretation of set inclusion to construct the second-level framework, **[3]** and **[10]**. Because the hypothesis refers to negative information about B, that is, $C\backslash B \subseteq C\backslash A$, our prover might think of doing a subproof by contradiction, and hence, suppose **[4]**, $x \notin B$. At this point, our prover would explore where this leads. Then use the operable interpretation of $x \in C \cap A$ to get $x \in$ C. This together with the operable interpretation of set difference gives **[5]**, $x \in C\backslash B$. Using the hypothesis that $C\backslash B \subseteq C\backslash A$ and the operable interpretation of set inclusion and modus ponens, gives **[6]**, $x \in C\backslash A$. Then the operable interpretation of set difference gives **[7]**, $x \notin A$, which is the contradiction pointed out in line **[8]**. Thus, it is legitimate to write **[9]**, $x \in B$. This finishes the proof, as lines **[10]** and **[11]** were already written.

The above proof is not particularly unusual or difficult to read. This includes the passage consisting of **[5]**, **[6]**, **[7]**, and **[8]**, even though no explicit warrants are provided. However, the explanation of it in the above commentary seems more difficult to understand. We suggest that this is because many readers have automated some of the actions involved. This illustrates the value of automation mentioned in the fourth paragraph of the section titled, "Theoretical Perspective".

Sample Incorrect Student Proof Attempt 3. Next we consider a student proof attempt of the same theorem.

Proof:
[1] Let A, B, and C be sets. **[2]** Suppose $C\backslash B \subseteq C\backslash A$. **[3]** Let $x \in C \cap A$. **[4]** Then $x \in C$ and $x \in A$.
[5] Since $C\backslash B \subseteq C\backslash A$, and $x \in C$ **[6]** but $x \notin B$, **[7]** $x \in C\backslash A$.

This student did not write the entire first-level framework, but started in the right place with the hypotheses. lines **[1]** and **[2]**. The student did not attempt a proof of $x \in$ B by contradiction despite being in a situation where he/she could not prove that $x \in B$ directly. Instead, the student seemingly began a direct proof in line **[3]**, taking an element of $C \cap A$ and unpacking what that meant **[4]**. After that, the student seemingly tried to use the hypothesis **[5]** and the fact that x is in C. It is not clear how the deduction, **[6]**, follows from what precedes it. It is possible that the student lost his/her tain of thought, and thought wrongly that he/she knew x $\in$ C\B. Also, there is no indication the student was starting a proof by contradiction. This leaves no reasonable way to conclude **[7]**.

It was not helpful that the student did not write a full proof framework. If he had, and noticed the he did not know how to continue with a direct proof, he might have seen how to start a proof by contradiction. We also note that the student drew two Venn diagrams, one in which both $C \subseteq A$ and $B \subseteq A$; we conjecture this was also not helpful.

Sample Correct Proof 4. Finally we consider a sample correct proof and an incorect student proof attempt of the following: *Theorem. Let X, Y, C, and D be sets and f:X→Y be a function. If* $C \subseteq D$ *and* $D \subseteq Y$*, then* $f^{-1}(C) \subseteq f^{-1}(D)$.

Proof:
[1] Let X, Y, C, and D be sets and $f{:}X \to Y$ be a function.
[2] Suppose $C \subseteq D$ and $D \subseteq Y$.
[3] Let x $\in X$. Suppose $x \in f^{-1}(C)$, **[4]** so that $f(x) \in C$.
[5] Then $f(x) \in D$, **[6]** which means $x \in f^{-1}(D)$.
[7] Therefore $f^{-1}(C) \subseteq f^{-1}(D)$. QED.

Our idealized prover would first write the first-level proof framework **[1]**, **[2]**, and **[7]**. By unpacking the conclusion **[7]**, our prover would know that he/she needed to start with an element of $f^{-1}(C)$, which is line **[3]**, and show $x \in f^{-1}(D)$, which is **[6]**, completing the seond-level framework. The operable interpretation of the definition of $f^{-1}(C)$ and yields **[4]**. Using the fact that $C \subseteq D$ our prover would get line **[5]**. Then applying the operable definition of $f^{-1}(D)$,gives **[6]**, and the theorem is proved.

Sample Incorrect Student Proof Attempt 4. Next we consider a student proof attempt of the same theorem, along with the student's scratchwork. In his/her work there was a large blank space between lines **[6]** and **[7]**.

Proof:
[1] Let X, Y, C, and D be sets and $f{:}X \to Y$ be a function.
[2] Suppose $C \subseteq D \subseteq Y$. **[3]** Suppose $y \in C$, **[4]** then $y \in D$ and $y \in Y$.
[5] Since f is a function, there is an $x \in X$ so that $(x,y) \in f$.
[6] Suppose $x \in f^{-1}(C)$, then

[7] Then $x \in f^{-1}(D)$.
[8] Therefore $f^{-1}(C) \subseteq f^{-1}(D)$.

Scratchwork:
Function: Then $f(D) = \{y \mid$ there is an element $d \in D$ so that $f(d) = y\}$

The student has written lines **[1]**, **[2]**, **[7]**, and **[8]** just as in the correct proof. So the student has written most of a proof framework, and to complete it, he should have written line **[6]** immediately after line **[2].** The student then introduces **[3]** $y \in C$ and legitimately concludes **[4]** y $\in D$, along with the extraneous fact that y $\in Y$. He/She then, but irrelevantly, states in line **[5]** that there is an element $x \in X$ so that $(x,y) \in f$. This assumes incorrectly, but irrelevantly that f is onto. Indeed, lines **[3]**, **[4]**, and **[5]** are not helpful. Then in line **[6**], it seems that the student begins again with the correct assumption, which had it been done earlier, would have produces a proof framework. Apparently he/she could not figure out how to get to line **[7]**, as indicate by the blank space. Although this proof was part of an in-class exam, the students were allowed access to all of their course notes. A beneficial action the student did not take would have been to write the operable interpretation of **[8]** into the scratchwork. Instead, it contains the definition of $f(D)$.

Summarizing, the student who wrote the above "proof" took a number of detrimental actions that *should not* have taken been and did *not* take a number of beneficial actions which that *should have* been taken.

Teaching Implications and Future Research

Having isolated and illustrated a few proving difficulties that our students, and probably many others, very often have, we can suggest some teaching interventions that might alleviate these difficulties. What form these interventions might take and how one might gauge their effectiveness, is a matter for future research. Because what might be done, and how to do it and gauge its effectiveness are closely intertwined we discuss them together.

Perhaps a good place to start explicit teaching is with proof frameworks, described in detail in the theoretical perspective. As noted above in the sample proofs, a number of difficulties seem to be traceable to not writing part or all of proof frameworks. Also, the writing of a proof framework can be decomposed into parts that can be taught separately. Perhaps an intervention might begin by thoroughly teaching students how to write one kind of common proof framework. After that the others could probably be learned quickly. We have found that students tend to resist writing full proof frameworks. We think this is because it involves writing in a way that is not "from the top down". In most of their past experience, texts were read and written from the top down. There should be enough practice for students, not only to understand what they are doing, but also to form a habit of consistently writing proof frameworks. That is, they should overcome their, possibly nonconscious, reluctance. Also, it would be good if the entire process of writing proof frameworks became automated. To accomplish this in a reasonable amount of time, it is probably better to ask students to practice constructing only a proof framework, not the entire proof, for each practice problem.

To gauge whether such an intervention has succeeded, one might interview students towards the end of the course, asking them to construct a few relatively easy proofs, and observe them to see if they wrote proof frameworks. One might also analyze examination proofs for difficulties that might be traceable to not having written a proof framework.

We turn now to another difficulty that occurs fairly often, namely, not correctly using definitions. Here we suggest the idea of operable interpretations, described in the theoretical perspective, would be good to explicitly teach. In transition-to-proof courses, there are likely to be many definitions. It would be useful if students knew their operable interpretation in an automated way. It would also be useful for students to eventually develop the ability to autonomously produce operable interpretations of formal definitions for themselves. To arrange both of these, one might consider occasional brief small group discussions developing operable interpretations for definitions about to be used in a course. At the end of the group discussions, a teacher might certify which interpretations would be accepted for the course. One might also consider very brief short-answer quizzes on collections of operable interpretations. Another kind of quiz might consist of a few fragments of proofs that students could extend a little using operable definitions. Again, to gauge the usefulness of such activity, one might want to draw on quizzes, interviews, and an analysis of some examination questions.

In the next iteration of the course, we hope to implement the above interventions and investigate their effectiveness.

References

Baddeley, A. (2000). Short-term and working memory. In E. Tulving & F. I. M.Craik (Eds.), *The Oxford Handbook of Memory* (pp.77-92). Oxford: Oxford University Press.

Baker, D., & Campbell, C. (2004). Fostering the development of mathematical thinking: Observations from a proofs course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 14*(4), 345-353.

Bargh, J. A. (1997). The automaticity of everyday life. In. R. S. Wyer, Jr. (Ed.), *The automaticity of everyday life* (pp. 1-61). Mahwah, NJ: Lawrence Erlbaum Associates.

Bargh, J. A. (2014). Our unconscious mind. *Scientific American*, *310*(1), 30-37.

Coppin, C. A., Mahavier, W. T., May, E. L., & Parker, G. E. (2009). *The Moore Method: A pathway to learner-centered instruction* (MAA Notes No. 75). Washington, DC: MAA.

Mahavier, W. S. (1999). What is the Moore Method? PRIMUS, 9, 339-354.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*, 249-266.

Norton, A., & D'Ambrosio, B. (2008). ZPC and ZPD: Zones of teaching and learning. *Journal for Research in Mathematics Education, 39*(3), 220-246.

Ranganath, C., Libby, K. A., & Wong, L. (2012). Human learning and memory. In K. Frankish & W. M. Ramsey (Eds.), *The Cambridge Handbook of Cognitive Science*. Cambridge: Cambridge University Press.

Selden, A., McKee, K., & Selden, J. (2010). Affect, behavioural schemas and the proving process, *International Journal of Mathematical Education in Science and Technology, 41*(2), 199-215.

Selden A., & Selden, J. (2014). The roles of behavioral schemas, persistence, and self-efficacy in proof construction. In B. Ubuz, C. Hasar, & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 248-255). Ankara, Turkey: Middle East Technical University.

Selden, A., & Selden, J. (2013a). The genre of proof. In M. N. Fried & T. Dreyfus (Eds.), *Mathematics and Mathematics Education: Searching for Common Ground*. New York: Springer, Advances in Mathematics Education series, 2013. Also, available as Tennessee Technological University Mathematics Department Technical Report No. 2013-1, which can be downloaded from http://www.tntech.edu/math/reports/.

Selden, A., & Selden, J. (2013b). Proof and Problem Solving at University Level. *The Mathematics Enthusiast*, Special Issue: International Perspectives on Problem Solving Research in Mathematics Education (Manuel Santos-Trigo & Luis Moreno-Armella, Guest Editors),Vol. 10, 1&2 (January 2013), 303-334.

Selden A., & Selden, J. (1987). Errors and misconceptions in college level theorem proving. In J. D. Novak (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics* Vol. III (pp. 457-470), Ithaca, NY: Cornell University.

Selden, J., & Selden, A. (2009). Teaching proving by coordinating aspects of proofs with students' abilities. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and Learning Proof Across Grades: A K-16 Perspective* (pp. 339-354). New York/Washington, DC: Routledge/National Council of Teachers of Mathematics.

Selden, J., & Selden, A. (2008). Consciousness in enacting procedural knowledge. *Proceedings of the 11th Annual Conference on Research on Mathematics Undergraduate Mathematics Education.* From http://rume.org/crume2008/Proceedings/Selden_Procedural%20LONG.pdf.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*, 123-151.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge, *Educational Studies in Mathematics, 48*, 101-119.

TECHNOLOGY AND ALGEBRA IN SECONDARY MATHEMATICS TEACHER PREPARATION PROGRAMS

Eryn M. Stehr and Lynette D. Guzman
Michigan State University

Most recently, the Conference Board of the Mathematical Sciences has advocated for incorporating technology in secondary mathematics classrooms. Colleges and universities across the United States are incorporating technology to varying degrees into their mathematics teacher preparation programs. This study examines preservice secondary mathematics teachers' opportunities to expand their knowledge of algebra through using technology and to learn how to incorporate technology when teaching algebra in mathematics classrooms. We explore the research question: What opportunities do secondary mathematics teacher preparation programs provide for PSTs to encounter technologies in learning algebra and learning to teach algebra? We examine data from a pilot study of three Midwestern teacher preparation programs conducted by the Preparing to Teach Algebra (PTA) project. Our data suggest that not all secondary mathematics teacher preparation programs integrate experiences with technology across mathematics courses, and that mathematics courses may provide few experiences with technology to PSTs beyond strictly computational.

Key words: Algebra and Algebraic Thinking, Technology, Preservice Teacher Education, High School Education

This study explores opportunities provided by secondary mathematics teacher preparation programs for preservice teachers (PSTs) to expand their knowledge of algebra through the use of technology and to learn how to incorporate technology when they teach algebra. We explore the following research question: *What opportunities do secondary mathematics teacher preparation programs provide for PSTs to encounter technologies in learning algebra and learning to teach algebra?* These opportunities might include observing, using, or learning about a variety of algebra-appropriate technologies, as well as thinking critically about technology use. In this study, we define technology narrowly as electronic tools and software. This study will not focus on physical tools such as manipulatives, chalkboards, or dry erase boards, although we acknowledge that these tools are also important technologies that can be useful for teaching and learning mathematics. Our stance is not to claim technology is necessarily useful or not useful; however, we primarily draw upon recommendations for mathematics teacher education from recent reports and documents published in the United States, which encourage critical choice and strategic use of technological tools through learning opportunities. The development of internal frameworks that support critical choice and strategic use of technological tools has been recommended for some time as an important part of teacher preparation programs by the Conference Board of Mathematical Sciences (CBMS, 2001; 2012) and also by national accreditation agencies (InTASC: CCSSO, 1995; NCATE: NCTM, 2012). From this perspective, we assume PSTs' opportunities to learn with technology are important components in their teacher preparation programs as they develop their professional teaching skills.

Context

Technology use in K-12 education has become practically universal in the past few decades. Many scholars suggest that use of technological tools in the classroom could contribute to reducing inequities in education and support subject-matter learning for a range of diverse students. For example, Pomerantz (1997) argued: "....Calculators serve as an equalizer in mathematics education" (p. 5). Technology use, however, has led to a so-called digital divide (Reich, Murnane, & Willett, 2012). Attewell and Gates (2001) described the digital divide as two-fold: a division of access and of use. Federal funding has mitigated issues of access; however, there is a growing recognition of disparity in technology use in schools (Attewell & Gates, 2001). Thus, a focus shifts from supplying schools with technology to considering the highly effective ways in which technology can be (but is not usually) used.

Both secondary mathematics content standards and teacher preparation standards have emphasized the importance of developing PSTs' abilities to critically choose and use educational technologies. Standards developed for teacher preparation program accreditation agencies, such as National Council for Accreditation of Teacher Education (NCATE: NCTM, 2012) and Interstate Teacher Assessment and Support Consortium (InTASC: CCSSO, 1995), recommended that PSTs develop the abilities to critically evaluate and strategically use technology. In addition, the Conference Board of the Mathematical Sciences (CBMS) emphasized the importance of PSTs' preparation to use technology in *Mathematics Education of Teachers II* (CBMS, 2012). In particular, CBMS (2012) recommends that PSTs should have multiple opportunities to engage with technologies in their own learning experiences and also develop the capacity to engage with technologies in teaching students in their mathematics classroom.

Algebra plays a prominent role in mathematics education reform efforts because it is valued as an important subject in mathematics. In terms of equity issues related to mathematics education, algebra has long been considered a gatekeeper for post-secondary education opportunities (e.g., Moses, Kamii, Swap, & Howard, 1989). Particularly in the United States, preparing future secondary mathematics teachers to teach algebra has gained importance as more states include algebra as a high school graduation requirement (Teuscher, Dingman, Nevels, & Reys, 2008). Consideration of state education websites verifies that at least 38 states currently include mathematics courses with algebra as a necessary high school graduation requirement. Algebra is also being offered earlier in some states. In 1990, only 16% of all eighth-graders were enrolled in algebra, and this percentage increased to 31% by 2007 (Loveless, 2008). The emphasis of algebra in mathematics education, along with increasing use of technology in the classroom, highlights the need to support future mathematics teachers in learning algebra with technology and learning to teach algebra with technology.

To use technology effectively to support the teaching of algebra, CBMS (2012) argued that experience with technology "should be integrated across the entire spectrum of undergraduate mathematics" (pp. 56-57) and PSTs should have opportunities to see teaching with technology modeled in their own mathematics coursework (CBMS, 2012). PSTs need to become familiar with a variety of technological tools used in a variety of ways, including computational tools, problem-solving tools, and tools for exploring mathematical ideas (CBMS, 2001; 2012; NCATE, 2012; InTASC, 1995). Naturally, we would expect to see a variety of technology use in mathematics courses as instructors model strategic choice of technology in the classroom.

Method

This study is part of a larger mixed-methods study, *Preparing to Teach Algebra (PTA),* which is exploring opportunities provided by secondary mathematics teacher preparation programs to learn algebra, to learn to teach algebra, to learn about issues in achieving equity in algebra learning, and to learn about algebra, functions, and modeling standards and mathematical practices as described in the *Common Core State Standards in Mathematics (CCSSM)*. The *PTA* project consists of a national survey of secondary mathematics teacher preparation programs and case studies of five universities. This paper focuses more narrowly on opportunities provided to PSTs to encounter technology in learning algebra and learning to teach algebra. This paper reports on results from a qualitative analysis based on data gathered from three university teacher preparation programs during the pilot study of the *PTA* project.

In the pilot study, the *PTA* project chose three secondary mathematics teacher preparation programs as a sample of convenience. University A is a medium-sized university based on enrollment with Carnegie classification of RU/H (Research University with high research activity). Universities B and C are large-sized universities based on enrollment, both with Carnegie classification RU/VH (Research University with very high research activity). The programs at Universities A and C are four-year programs, and the program at University B is a five-year program.

We compiled data by conducting five instructor interviews and one focus group interview at each site. Each instructor interview focused on a previously selected course that was chosen by the researchers for its potential to include opportunities for PSTs to learn algebra or learn to teach algebra, and also according to availability and course type. The only exceptions to this were at University B, where four mathematics education courses are taught as two year-long sequences. One instructor was interviewed about each year-long sequence and asked to focus on the first course of the year (the 1st and 3rd Secondary Math Methods courses, respectively). This focus on only part of the year-long course proved difficult, so we collected and analyzed instructional materials from all four courses. We treated them in the analysis as two year-long courses, except for data that we knew came from a particular semester.

At each university, we attempted to balance representation of course types by choosing two mathematics courses for all mathematics majors, one mathematics course designed specifically for PSTs, and two mathematics education courses, as shown in Table 1 below. Each course was required by the program and had been recently taught by the instructor that we interviewed. We collected corresponding University course descriptions and instructional materials from each instructor. Among other questions in each interview, we asked instructors which types of technologies they used in a particular course; we also analyzed their course materials.

Table 1. Chosen courses at each site.

Type of Course	University A	University B	University C
Mathematics	Linear Algebra	Linear Algebra	Differential Equations
Mathematics	Structure of Algebra	Analysis	Abstract Algebra
Mathematics Designed for PSTs	Secondary Math from an Advanced Viewpoint	Math Capstone	Seminar
Mathematics Education	Middle School Math Methods	1st and 2nd Secondary Math Methods*	Middle School Math Methods

Mathematics Education	Secondary Math Methods	3rd and 4th Secondary Math Methods*	Secondary Math Methods

*Note: At University B, 1st and 2nd Secondary Math Methods is a one-year sequence; 3rd and 4th Secondary Math Methods is a one-year sequence. We treated each one-year sequence as a whole course.

We conducted focus group interviews with three or four students who had completed, or had almost completed, their student teaching requirement in each program. We asked PSTs to elaborate on their required and shared experiences with algebra across the entire program, including the five required courses listed above in Table 1. They confirmed a list of program requirements and identified which required courses incorporated technology in learning algebra or learning to teach algebra.

One important note is that our unit of study is the teacher preparation program as a whole. We do not intend to evaluate or compare the programs in our study. Data gathered from PSTs is not necessarily representative of the same courses as data gathered from instructors because PSTs almost certainly did not take the specific enactment of a course described by our interviewed instructors.

Because this study uses pilot data from a larger study, one limitation is that instructor interviews were restricted to five courses at each site and that these courses were not representative of an entire teacher preparation program. Additionally, we chose courses based on their likelihood to contain algebraic content and not specifically for a focus on technology. As a result, we missed data on other courses that provided additional opportunities for PSTs to experience technology in secondary mathematics. To balance this limitation, we used information from focus groups and course descriptions obtained from school websites to create an outline sketch of technology use across each program.

Results

We first give a brief report of the results obtained from our analysis of examples of technology use in algebra teaching and learning. Because explicit examples of algebra and technology use were limited due to the type of data we collected, we also gathered statements from instructors regarding why they did or why they did not choose to use technology. We follow the report of technology at each university with a presentation of the themes we saw in the instructors' explanations of why they would or would not use (or allow) technology in their courses, whether used by students or the instructor.

University Technology Use

To create an outline sketch of technology use at each university, we first identified and gathered specific examples of technology use in algebra from instructor interviews, focus group interviews, or from the instructional materials. We analyzed each example according to five characteristics of experiences: activity type, types of technology use, algebraic topics, type of technology, and whether PSTs had the opportunity to think critically about choice and use of technology.

Across all universities, we found 28 explicit examples of technology use in algebra. This count excludes numerous examples in a Differential Equations course at University C, which involved a computer lab component. Of the 28 examples of technology use in algebra, eight come from mathematics content courses and 20 from mathematics education courses. Table 2 shows a descriptive list of algebraic topics and in which courses examples were found (M for Mathematics courses, including mathematics courses designed for PSTs, and ME for Mathematics Education courses).

Table 2. Algebra topics using technology identified per university and by mathematics or mathematics education courses.

Algebraic Topics	Univ. A	Univ. B	Univ. C
Generalizing Patterns		ME	
Maximum Area Problem		ME	
Ratios and Proportion		ME	ME
Modeling with Equations			ME
Functions and Multiple Representations	ME	ME	MfT
Linear Functions (e.g., families, slopes)		ME	ME
Systems of Linear Equations	M, MfT		
Parametric Equations		ME	
Logarithmic Functions	ME		
Matrices	M, MfT		
Topics from Differential Equations			M
Modular Arithmetic	M		
Extensions on Rational Numbers	MfT		

University A: Overview of technology use. University A requires twelve mathematics courses and four mathematics education courses. The Mathematics Department offers the mathematics education courses. Students are required to take a mathematics education course titled "Teaching Secondary Mathematics with Technology." In this study, the technology mathematics education course was not a focus of an interview, but other mathematics education instructors described some of its content. For example, an instructor noted that PSTs used GeoGebra and Google SketchUp in the Teaching Secondary Mathematics with Technology course. Another instructor said of the program as a whole, "We think [technology is] crucial," which is a statement supported by both the number of mathematics courses that include technology use in their course descriptions (six of twelve) and the program requirement of one mathematics education course focused on teaching with technology. Overall, student and instructor responses indicated technology was used in mathematics courses primarily as a computational tool, while mathematics education courses supported a greater variety of types of uses of technology, including some critical evaluation of technology for algebra teaching.

One example of technology use with algebra at University A includes an activity in the Teaching Secondary Mathematics with Technology course in which PSTs "investigate graphing utilities and think about what are the features of graphing utilities that would … make one more desirable than another." In this assignment, neither instructors nor PSTs necessarily used the technology, but students thought critically about multiple possible uses and types of graphing utilities. In a second example, the Secondary Math Methods instructor described discussions about preparing unit plans, particularly regarding the introduction of systems of linear equations or logarithmic functions, and determining what advantages or disadvantages come with technology use for those topics. Finally, a third example of technology use with algebra comes from the Structure of Algebra course instructor, who stated, "when we talk about cryptography I'll bring in Mathematica... if you want to do RSA cryptography in any sort of realistic way, you want to use... you know, RSA relies on a number that's a product of large primes. …So you're doing … arithmetic mod some huge number."

University B: Overview of technology use. University B requires eight mathematics courses and four mathematics education courses. The College of Education offers the mathematics education courses. PSTs in the focus group marked some use of computer software in Calculus III, geometry courses, and statistics courses, as well as multiple technologies in the four mathematics education courses. One student in the focus group stated, "…tools for me is by far the biggest weakness… even when we did use them it was pretty rare."

Overall, we found few examples of technology use in mathematics courses at University B. Instructors of Linear Algebra, Analysis, and the Capstone course stated that they did not use technology in class, except occasionally to check a calculation. The Linear Algebra instructor explained, "I don't think it is a good idea to use calculator or computer software… you want them to do it by hand." The mathematics education courses used multiple instructional and mathematical technologies to support algebra topics and some critical evaluation of technology. Specific mathematical technologies included GeoGebra, spreadsheets, graphing calculators, and the occasional use of Geometer's Sketchpad. One instructor explained that he chose to use technology because "…[the PSTs] see things mathematically they didn't see before and it helps them see the value of engaging in those sorts of tasks with their students…."

Additionally, student and instructor responses indicated that few mathematics courses used technology in learning opportunities, while mathematics education courses integrated a variety of technologies to support PSTs' teaching and learning of mathematics as well as PSTs critical evaluation of technological tools. One example of technology use in algebra was in the 2nd Secondary Methods course, the instructor introduced students to the "Ships in the Fog" task (based on the crash of the Stockholm and Andrea Daria) through a newsreel video of the wreck, solving the problem three ways (the worksheet calls for graphing calculator use), discussing the task on the Wiki, and then reading the "Ships in the Fog" case. PSTs used multiple technological tools as they watched the video clip, solved the problem with graphing calculators in multiple ways, and discussed their solution strategies on the Wiki site.

University C: Overview of technology use. University C requires twelve mathematics courses and two mathematics education courses. The College of Education offers the mathematics education courses. An Educational Technology course is a required general education course in the program but does not focus on mathematics. The mathematics department at University C implements a policy that does not allow graphing calculators on mathematics final exams. PSTs indicated four mathematics courses in which they used computer software or clickers, although they did not acknowledge technology use in mathematics education courses. PSTs stated that they did not learn to use certain technologies despite needing them later in field instruction. In her notes from the focus group, one PST wrote along the list of program courses, "no graphing or non-graphing calculator allowed." The Abstract Algebra and Differential Equations instructors indicated rare use of technology in lectures; however, the Differential Equations course included a computer lab component using MatLab.

Although the mathematics education course instructors at University C did not emphasize technology when talking about their instruction, one assignment in the Secondary Mathematics Methods course did require students to revise a previously written lesson plan to "include technologies that *enhance the teaching and learning of mathematics*," and to discuss their rationale for inclusion. Through this assignment, PSTs had explicit opportunities to consider how technology use in the classroom may enhance students' opportunities to learn mathematics. Additionally, PSTs received feedback through a draft, peer review, and revising process over the course of a semester for this assignment.

Overall, student and instructor responses at University C indicated several mathematics courses used technology, while mathematics education courses supported critical evaluation through choice and justification of technology for mathematics teaching, although the courses themselves did not integrate technology use. One specific example of technology use in algebra at University C was a lab activity in the Differential Equations course that focused on the existence and uniqueness theorems. Through creating graphs of several solution sets, students were asked to explain solutions in terms of the theorems, explain why an equation did not satisfy hypotheses of the theorems, make claims and explain why a solution must exist and be unique, investigate limitations of the theorems, investigate why and how results could be wrong answers, and make and investigate claims about boundedness of solutions. In this way, technology was used to help students bridge the gap between abstract theorems and applications. Students experienced engagement in mathematical practices such as claiming, explaining, and investigating.

Instructor Rationales For or Against Technology Use

Both mathematics and mathematics education instructors described rationales for using technology or supporting students' use of technology. Instructors of both types of courses also described rationales for not using technology or actively discouraging students from using technology. We present these rationales as falling into four categories: practical concerns, beliefs that technology impedes mathematical learning, beliefs that technology enhances mathematical learning, and acknowledgements that the decision is complicated.

Practical concerns. Several instructors reported practical concerns that kept them from using technology, referencing the appropriateness of technology in particular courses as well as issues of access, time, and support. For example, the Abstract Algebra instructor from University C argued that technology was not appropriate in such a course, stating, "No calculators, spreadsheets, no SmartBoard… really low tech. It's abstract for a reason." A number of other instructors we interviewed expressed a similar viewpoint about technology not being helpful in pursuing the goals of their courses.

At both Universities A and B, mathematics education instructors stated that they were unable to use certain technologies because scheduling time to integrate technology into instruction was difficult and the department simply did not have the money. For example, the University A Secondary Math Methods instructor said, "…we don't have money in our department to buy [a SmartBoard or clickers]. So we don't have those and our students … need to know those things." At University B, mathematics instructors indicated a need for additional time and support to prepare for technology use. For example, the Analysis instructor described wishing she could use a particular computer simulation that could help students see the differences between convergence and uniform convergence of functions. She said she could see it in her mind and would like to be able to show her mental image to the students. She drew the example on the chalkboard and described what she would like a computer simulation to demonstrate, saying "I think one day maybe we should have a nice computer-simulated programs that make you see the difference. It would be really nice."

Impeding learning. Instructors of both mathematics and mathematics education courses also described their beliefs that technology can impede learning. For example, the University B Analysis instructor said she believed that "…because of the calculator and all these technologies [people] don't … develop their memory. But then you are asking them to develop their memory on something that is harder than adding or subtracting, you know?" The University A Linear Algebra instructor also held a perspective of technology potentially impeding student learning by stating, "…there are some computationally intensive problems in the book so in those cases I expect them to use technology… But I also want them to know the concepts involved so sometimes you know I make a point to tell them that they shouldn't use technology…" The University A Secondary Math Methods instructor similarly said that "at a college level we're now quite concerned because we have students who can't multiply… We have a huge problem in Calculus with kids who've come out of high school with A's and B's. And so we have students who can't multiply. We have students who can't reduce a fraction… I think because they have always had a calculator, you know. There are students who can't tell you what the graph of $y = x$ looks like. They could produce it on the graphing calculator, but to be able to think about what $y = x$ and $y = x^2$ looks like – they can't do it without a machine… So we are actually moving to not using technology." Instructors' perspectives towards how technology may influence students' learning of mathematics seemed to influence the opportunities they provided PSTs to encounter technologies in their courses.

Enhancing learning. Instructors of mathematics education courses described ways that technology can enhance learning by making the abstract more tangible, allowing different perspectives, and supporting PSTs' development of conceptualizations of mathematics. For example, the University C Middle School Math Methods instructor said, "I strongly encourage them to use [technological tools] as much as they can because I think that there's different types of learners and that sometimes a hands-on or a computer simulation, getting up to a SmartBoard and drawing out your thinking in some way, …they can bring some of these more abstract things to make them more tangible for students. I wouldn't say I've done the best job of exemplifying this in this class, but I do encourage it." The University B 1st and 2nd Secondary Math Methods instructor said, "I think it - it enhances their ability to, you know, model situations and … gives them a way to see the problem from a different perspective, one that they might not be familiar with such as a visual representation of a binomial or a trinomial and so it helps them kind of understand it from a learner's perspective and also gives them ways to think about how to instruct students in multiple ways so how they might teach with tools that can be accessible to learners." The University B 3rd and 4th Secondary Math Methods instructor said, "All of these tools represent ways to represent and conceptualize mathematical ideas that go beyond the symbolic. They're important tools to really develop a conceptual understanding of mathematics. Moreover, it's critical that our students are prepared to use these same tools with their own students in the classroom to foster the same sorts of understandings."

Strategic technology use is not straightforward. Both mathematics and mathematics education instructors argued that technology is not appropriate in every course and that instructional consequences should be considered.

For example, introducing unfamiliar software can shift the focus of a course from learning new mathematics to learning a new software. The University C Differential Equations instructor described his thinking when deciding whether or not to use technology in a particular course, explaining that in other courses he taught he did choose to use technology. He explained that he had not introduced technology into his Differential Equations course, saying, "[Calculator and graphing activities] are not really related to the basic themes and goals that I think are best in this course and as I said before I don't want to blur the focus." He went on to describe the shift in focus that would be required to teach students to use MATLAB in addition to the course content, saying, "The course is pretty close-packed and I don't have time to teach them MATLAB in this course… and the more technology you have in a course like this the less that there is for algebra."

Instructors of mathematics education courses also described their discussions with students about why technology needs to be used at times but should not be used at other times. The University A Secondary Math Methods instructor described her concerns about use of technology, saying, "...there are times where instructionally it may be not the best thing to always use technology and so making that kind of judicious choice is something we talk about as well." The University B 1st and 2nd Secondary Math Methods instructor similarly described her discussions with students about when technology should be used, saying, "... you don't just use a tool or technology just because it's going to be fun; but you really have to think about - What does this particular tool or technology afford me in terms of students' understanding the content?" She also explained that her class discussed how thinking about use of technology can be a complicated balance between instructor's time, instructional time, and benefits, saying "they come out of the course realizing that it's a lot of work - it takes time and so I have to be very strategic in how I use tools and technology" and that they reflect on their use of technology and when technology really provides something different than other tools, saying "…sometimes when we've used technology it didn't really offer us any more than if we had just drawn [on] a piece of paper…"

Discussion

Contrary to CBMS (2012) recommendations, our data suggest that not all secondary mathematics teacher preparation programs integrate experiences with technology across mathematics courses. We found that mathematics education courses integrate technology into instruction and learning more commonly, and with more variety in types of use, than mathematics courses. Even in mathematics courses that use technology, our data suggest that PSTs have fewer opportunities to see and use a variety of technological tools and that PSTs are more likely to see or use technologies only as computational tools. With respect to specific experiences using technology in learning and learning to teach algebraic topics, according to our data, mathematics education courses provide the bulk of these experiences.

We heard concerns from both mathematics and mathematics education instructors that technology would impede PSTs' learning. Some mathematics education instructors argued, to the contrary, that use of technology enabled PSTs to increase their understanding of algebra topics in ways that were not possible otherwise. One explanation of this difference in instructors' viewpoints might lie in whether instructors used technology only as a practical expedient.

We also heard that mathematical and instructional consequences should be considered before using technology. For example, instructors described the importance of considering the amount of time needed by an instructor to implement technology outside of class as well as within. Creating PowerPoint slides, designing applets, or developing other instructional materials and activities that use technology can be time-consuming for the instructor. Teaching students to use a technological tool or allowing students to become familiar with technology may divert time within a course that could otherwise be spent on the subject matter. Using technology may also change the way students conceptualize a mathematical concept or the way aspects of a topic are emphasized, whether intentionally or not. For example, using mathematical software or graphing calculators to solve systems of linear equations may not allow students to learn row operations, but could be used to highlight patterns in solutions of systems of linear equations instead. Using mathematical software to graph a quadratic function may impede students' learning, but using mathematical software to dynamically explore the impact of the parameters of a quadratic function may help students understand those parameters in a different way than they could by using paper-and-pencil. These examples illustrate that different ways of understanding mathematics and different emphases can be supported with different tools, whether paper-and-pencil or mathematical software. Instructors we spoke to recommended that mathematical consequences such as these should also be considered when deciding whether technology should be used or not.

We described some of the careful thought that instructors of both mathematics and mathematics education courses shared with us regarding how they have chosen to use or not to use technology in their courses. As described in the university overviews, we found that at every university, PSTs had opportunities to think critically about use of technology in both in-class discussions as well as assignments. One assignment that we saw in some form at each university involved either critically comparing graphing utilities as at University A, or incorporating some technology into a previously written lesson plan at Universities B and C. Based on the assignment rubric, the technology assignment at University C emphasized presenting a strong argument that technology provided something new and valuable that would not otherwise be provided. Mathematics education instructors at each university also described discussions with PSTs about the possible negative consequences of using technology in mathematics in addition to discussions about affordances. These opportunities to discuss and create arguments for technology will allow PSTs to develop the critical framework recommended by *METI* and *METII* (CBMS, 2001; 2012).

Limitations. We acknowledge some important limitations of our data and analysis. As described in the Method section, our data was gathered in the pilot phase of the *Preparing to Teach Algebra* project. Some limitation of this study are being addressed in the main phase of the *PTA* project. During the pilot phase, despite balancing the types of the courses, the courses were a sample of convenience and not necessarily representative of algebraic content in the program as a whole. The courses were also chosen for algebra content, and not necessarily for technology use, and we know we missed at least one course (Teaching Secondary Mathematics with Technology) and possibly others. For the main study phase of the *PTA* project, courses were chosen more carefully to be representative of each program, while still being chosen for their potential to address our research questions.

In addition, the collection of instructional materials depended on the interviewer and instructor. We collected at least a syllabus from each course, but the amount of instructional materials varied greatly from course to course. Some course instructors gave nothing but the syllabus while others gave a sampling of exams or quizzes. The University B Secondary Math Methods and University C Differential Equations courses provided the bulk of the instructional material that was analyzed, possibly because the instructors posted the materials on a website and gave the researchers access. The *PTA* project has been more systematically selective in the instructional materials gathered and analyzed for the main study phase.

Other limitations are inherent in the choice of data and necessarily limited scope of the research questions. Our data was gathered from interviews and written materials, without observing classroom practice. We also only interviewed one instructor per course, even though different instructors can teach the same course in very different ways. We have attempted to balance these possible limitations by gathering data from instructors and students, hoping these different perspective will give us a reasonable understanding of the program as a whole.

Conclusions

We end this paper by expressing our hopes that it will raise questions about when, where, and how technology should appear in mathematics and mathematics education courses. *METI* and *METII*, as well as accreditation documents, recommend that PSTs encounter technology use and evaluation of technology use in their teacher preparation programs across both mathematics and mathematics education courses (CBMS, 2001; 2012). These documents recommend that encounters with technology in mathematics courses should support PSTs mathematical learning and in mathematics education courses should include critical evaluation of technology use to support PSTs development of an internal framework that could support their later critical choice and strategic use of technology in their own teaching. From this set of data, we think an emphasis should be on turning technology use in mathematics courses to be more than just computational use. Mathematics courses, like the University C Differential Equations course and as suggested by the University B Analysis instructor, could use mathematical software to help students make connections between abstract theory and applications by critically investigating and questioning the theory using dynamic representations and applied examples. We hope this paper challenges perspectives toward technology use in mathematics. Specifically, to challenge the bias that technology can only be useful in applied mathematics courses and that technology always impedes mathematical learning. We also hope that the practical difficulties faculty face, such as lack of time, support, or access, can be acknowledged and possible solutions explored.

Further research should be planned to investigate ways technology can be used more effectively in algebra to support future teachers' understanding of algebra as well as their abilities to choose and use technology more effectively in their own classrooms. Our research findings also point to additional questions that should be considered in when, where, and how PSTs should encounter technologies in opportunities to learn algebra and to learn to teach algebra as they develop their professional skills.

Endnote

This study comes from the *Preparing to Teach Algebra* project, a collaborative project between groups at Michigan State (PI: Sharon Senk) and Purdue (co-PIs: Yukiko Maeda and Jill Newton) Universities. This research is supported by the National Science Foundation grant DRL-1109256.

References

Attewell, P., & Gates, L. (2001). Comment: The First and Second Digital Divides. *Sociology of Education*, *74*(3), 252–259.

Conference Board of the Mathematical Sciences [CBMS]. (2001). *The Mathematical Education of Teachers*. Providence RI and Washington DC: American Mathematical Society and Mathematical Association of America.

Conference Board of the Mathematical Sciences [CBMS]. (2012). *The Mathematical Education of Teachers II*. Providence RI and Washington DC: American Mathematical Society and Mathematical Association of America.

Council of Chief State School Officers [CCSSO]. (2011). *Interstate Teacher Assessment and Support Consortium (InTASC) Model Core Teaching Standards: A Resource for State Dialogue*. Washington, DC: Author.

Loveless, T. (2008). *The misplaced math student: lost in eight-grade algebra*. The Brown Center Report on American Education.

Moses, R. P., Kamii, M., Swap, S. M., Howard, J. (1989). The Algebra Project: Organizing in the spirit of Ella. *Harvard Educational Review*, 59(4), 423-443.

National Council of Teachers of Mathematics (NCTM). (2012). *NCATE/NCTM Program standards: Programs for initial preparation of mathematics teachers.* Retrieved October 12, 2012, from http://www.nctm.org/uploadedFiles/Math_Standards/NCTM%20NCATE%20Standards%20Revision%20and%20Math%20Content%20-%20Secondary.pdf

Pomerantz, H. (1997). *The role of calculators in math education (Research compiled under the direction of Bert Waits)*. Columbus: The Ohio State University.

Reich, J., Murnane, R., & Willett, J. (2012). The State of Wiki Usage in U.S. K-12 Schools: Leveraging Web 2.0 Data Warehouses to Assess Quality and Equity in Online Learning Environments. *Educational Researcher*, *41*(1), 7–15. doi:10.3102/0013189X11427083
Teuscher, D., Dingman, S. W., Nevels, N. N., & Reys, B. J. (2008). Curriculum Standards, Course Requirements, and Mandated Assessments for High School Mathematics: A Status Report of State Policies*.* *NCSM Journal*, 49.

PERCEPTIONS IN ABSTRACT ALGEBRA: IDENTIFYING MAJOR CONCEPTS AND CONCEPT CONNECTIONS WITHIN ABSTRACT ALGEBRA

Ashley L. Suominen
University of Georgia

Abstract algebra is recognized as a highly problematic course for most undergraduate students. Despite these difficulties, most mathematicians and mathematics educators affirm its importance to undergraduate mathematical learning. The purpose of this research was to formulate a list of the important concepts in abstract algebra as perceived by graduate students in mathematics, understand how they define these concepts, and recognize any relationships or connections between these concepts perceived by the students. The theoretical perspective of concept images and concept definitions as described by Tall and Vinner (1981) and Vinner (1983) was used to investigate participants' understanding of abstract algebra concepts. Through an interview study, the students' perceptions were analyzed through the creation of concept maps. The results revealed the participants had great difficulty articulating their concept images and concept definitions. In addition, they had differing views of major concepts and relationships within the course.

Key words: Abstract algebra, Concept maps, Concept image, Concept definition, Connections

Introduction

It is widely acknowledged that abstract algebra is an essential part of undergraduate mathematical learning (e.g., Gallian, 1990; Hazzan, 1999; Selden and Selden, 1987), and yet it is also known for its high level of difficulty at the collegiate level. Many undergraduate and graduate students, including prospective teachers, struggle to grasp even the most fundamental concepts of this course (Dubinsky et al, 1994). For many of these students abstract algebra is the first time they experience mathematical abstraction and formal proof, and it is often the first course in which teachers expect students to "go beyond learning 'imitative behavior patterns' for mimicking the solution of a large number of variations on a small number of themes (problems)" (Dubinsky et al., 1994, p. 268). As a result, we can only expect abstract algebra students to really access the benefits of this course through the development of accurate mathematical meanings of the course concepts amidst the abstraction. This development will typically involve personally constructed concept images and concept definitions that can be used to understand the abstract theories and ideas. Despite the importance of abstract algebra and the known difficulties of the subject, little research has been devoted to these concept images and concept definitions. This article will highlight the constructed concept images and concept definitions of mathematics graduate students to explore their perceptions' of concept importance and concept connections within the course.

Theoretical Perspective

This study utilizes the theoretical perspective of concept images and concept definitions as described by Tall and Vinner (1981) and Vinner (1983) to investigate participants' understanding of abstract algebra concepts. In using this theoretical lens it is believed that the formation of a mathematical concept often involves the development of both a concept image and a concept definition. Tall and Vinner (1981) defined concept image as "the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes." For instance, students learning abstract algebra may construct mental pictures of specific algebraic structures in attempting to understand them, which may include a list of properties, class activities, or related previously constructed concept images. For instance, a student's concept image of a group might include concept images about associativity, closure, and identity and inverse elements as well as a visual picture of an equilateral triangle and its rotations. A concept definition is then the "verbal definition that accurately explains the concept in a non-circular way" (Vinner, 1983). Ideally a person's concept definition would be the verbal description of her personal reconstruction of the concept image, but often a person's concept definition does not align with the formally accepted mathematical definition. To put it simply, a personally constructed concept image may seem sensible to that person and yet have some discrepancies in the concept definition when compared to the accurate definition. As a result, the concept images and concept definitions constructed by undergraduate and graduate students learning abstract algebra may or may not correspond to those taught in the course or found in the textbook.

Ideally when a professor introduces the formal definition of a concept, any previously constructed images about or closely related to the concept should transform to include this definition. As a result, the concept image becomes more robust and the concept definition more accurate. However, this scenario does not occur as often as most professors would hope or expect. Rather, Vinner (1983) described two additional scenarios that can occur upon introducing a concept definition to previously constructed concept images. In the first scenario the formal concept definition has an influence on the student's concept image for a short while but eventually is replaced with a personally constructed concept definition that fits the current concept image. While aspects of the formal concept definition may be integrated into the concept image, the formal concept definition in its entirety is not. In the second scenario the student considers the taught concept definition and the constructed concept image as two separate entities in which the student utilizes in different ways (i.e. the student may provide the formal concept definition when asked to by the professor but may rely on the concept image all other times). Unfortunately, the latter scenario is often a result of students relying on rote memorization to define a concept rather than a personal reconstruction of their concept images. As a result, the concept definitions are typically forgotten when not actively mentally used.

In tertiary mathematics, however, students are routinely taught concept definitions when no previously constructed concept image exists and concept images of closely related concepts are not called upon. In this situation the taught formal concept definition forms a new concept image. Ideally the new concept image and taught concept definition begin to shape each other as both develop simultaneously. However, often times the concept image is entirely constructed by taught concept definition, so when students rely solely on rote memorization to define a concept, the concept definition is again forgotten as well as the concept image.

Literature Review

Despite the importance of abstract algebra, the known difficulties of the subject, and an increasing amount of research on teaching and learning collegiate mathematics, few studies concentrate solely on abstract algebra. Past research in abstract algebra can be classified into three categories: student learning (e.g., Asiala, Brown, Kleiman, & Mathews, 1998; Brown, DeVries, Dubinsky, & Thomas, 1997; Leron, Hazzan, & Zazkis, 1995), teaching methods (e.g., Asiala, Dubinsky, Mathews, Morics, & Oktaç, 1997; Freedman, 1983; Pedersen, 1972), and proof writing (e.g., Hart, 1994; Selden & Selden, 1987; Weber 2001).

Research on the teaching and learning of abstract algebra has indicated the conceptual understanding of undergraduate students in abstract algebra is less than satisfactory (e.g. Dubinsky, Dautermann, Leron, & Zazkis, 1994; Hazzan & Leron, 1996). Leron and Dubinsky (1995) declared that both professors and undergraduate students view the teaching of abstract algebra as a disaster. Thus, several researchers have introduced alternative teaching approaches. One of the earliest papers on teaching of abstract algebra was Pedersen (1972) in presenting a learning activity that involved 10 paper equilateral triangles in which students folded them in various ways to evolve the non-cyclic 6-group. This hands-on, discovery-learning activity encouraged students to develop mental concepts images and accurate concept definitions of the non-cyclic 6-group in a unique way. Similarly, Huetinck (1996) introduced the SNAP learning activity to introduction group theory in which students initially rotate and translate an equilateral triangle on an overhead transparency sheet to explore all possible orientations. Student then use a nine-peg 3x3 square array board with three rubber bands to discover patterns through various reorientations of the rubber bands. Additional researchers (e.g. Burn, 1996; Freudenthal, 1973; Larsen, 2004) suggested the introduction of group theory through permutations and symmetries of an equilateral triangle. Larsen and Lockwood (2013) then modified this activity by having students search for parity in the group of symmetries to present the concept of quotient group. Larsen (2004, 2013) also employed these and similar activities in his Realistic Mathematics Education (RME) theory research that stressed teaching group theory through examples. Cook (2012) then paralleled the work of Larsen in the teaching and learning of rings, fields, and integral domains. In these past studies researchers utilized activity-based learning to build upon students' intuition and past mathematical knowledge to develop mathematical meaning of abstract concepts. Ultimately these activities should elicit more robust concept images and more accurate concept definitions rather traditionally lecture-based approaches.

Not building upon past concept images is often a stumbling block to students learning abstract algebra. In his dissertation, Cook (2012) asserted the difficulty students experience in abstract algebra is due to the lack of established connections between undergraduate mathematics and school mathematics. He affirmed that abstract algebra students "do not build upon their elementary understandings of algebra, leaving them unable to communicate traces of any deep and unifying ideas that govern the subject" (p. xvi). Fennema and Franke (1992) supported this theory: "If teachers do not know how to translate those abstractions into a form that enables learners to relate the mathematics to what they already know, they will not learn with understanding" (p. 153). These conjectures imply that tertiary professors must be able to not only convey an abstract idea to students but also provide students the opportunity to build mathematical meaning by relating the new concept definitions to previously constructed concept images.

While several research studies have focused on the teaching and learning of abstract algebra, no research has been done on the explicit concept images and concept definitions students have developed about abstract algebra. Likewise, students' perceptions of concept importance and concept connections have not thoroughly been studied. Therefore, the accumulation of the current literature in addition to the lack of literature led to the following research questions:

1) What are the important concepts of an abstract algebra course as perceived by mathematics graduate students and how are these concepts defined?
2) What are the conceptual connections or relationships between these concepts that are perceived by mathematics graduate students?

Methodology

In this research study three participants (pseudonyms: Andrew, April, and Heather) were purposefully selected based on two criteria: participants were accepted into at least the master's level mathematics graduate program and participants recently (within a year or less) enrolled in the master's level abstract algebra course. While undergraduate students typically take abstract algebra, graduate students were specifically chosen to provide an additional level of expertise. Prior to conducting this research, April and Heather had taken three lecture-based abstract algebra courses—an introductory course as an undergraduate and a yearlong sequence of two courses as a graduate student—and Andrew was concurrently enrolled in the third semester of the yearlong sequence with an additional. During the time of the study Andrew was in his first year of the doctoral mathematics graduate program with a pure mathematics emphasis, April was in her second year of the same program with a pure mathematics emphasis, and Heather was in her second year of the doctoral mathematics education graduate program. Andrew and April both have previous undergraduate mathematics background in applied mathematics, whereas Heather's undergraduate experience focused more on mathematics education. These past and current experiences will ultimate influence the formed concept images of topics in abstract algebra.

This research employed a semi-structured interview protocol with both open-ended questions and a construction task (Patton, 2002; Taylor & Bogdan, 1984; Zazkis & Hazzan, 1999). Each interview was audio recorded and ran approximately 45-60 minutes in length in a private room to ensure confidentiality. After the interviews were complete the audio was transcribed within a week of the interview. Since the purpose of this research study is to gain insight into graduate students' perspectives of abstract algebra, one of the central foci of the interview was the creation of concept maps. These maps allowed the participants and researcher to visually understand described relationships between concepts. Novak and Cañas (2008) and Trochim (1989) largely contributed to the overall research design of this activity. First, each participant was given index cards (or post-it notes) and asked to write any important or key concepts of abstract algebra on a card (one per card). When he or she was finished with this task, the participant was asked to explain each concept. Next, participants were asked to visually represent any conceptual relationships between these topics by placing their concept cards on a sheet of poster board and drawing lines or arrows between concepts that have some type of relationship. After each participant completed a concept map, he or she was asked to explain why each line was drawn. Grounded theory was then utilized when analyzing the data. Once the interviews were conducted and transcribed, the transcripts were coded and analyzed thematically (Charmaz, 2000; Patton, 2002; Taylor & Bogdan, 1984) focusing on the students' responses to concept importance and concept connections. Each participant's concept image and concept definition for the discussed important concept were then formulated based on their general talk and created concept map.

Results

As to be expected, each of the participants had a differing concept image and concept definition of major abstract algebra concepts. When asked to identify these concepts, April and Heather equated the time spent in class to the importance of the concept. April stated, "I think that fields are very important because we spent a lot of time discussing the different properties of fields and the different types of fields… So I felt it was really important." Likewise, Heather repeatedly defined concept importance by the number of days the professor discussed it in class.

Andrew, on the other hand, relied on his perceived usefulness of a certain concept to determine major concepts. When asked to describe ring theory Andrew stated:

> It's like you encounter rings first from like the first time you encounter math to be like the real numbers. We actually use them in our real life and everything, so in a way like this concept of rings kind of formalizes our understanding of what everything actually means.

In addition, throughout the interview Andrew continually revisited useful applications for a group so that he felt comfortable identifying the concept as important. A complete summary of the perceived important concepts of each student and their overlapped important concepts is found in Figure 1. However, despite the varying concept images associated with concept importance, there were five identified concepts that were mentioned by all three students: groups, rings, fields, Galois theory, and isometries with geometric applications. In addition to these abstract algebra concepts, all three participants acknowledged the importance of learning about mathematical definitions, notion, and proof in abstract algebra. Andrew mentioned that in abstract algebra you "formalize everything, like in the practical in proper mathematical notation." April highlighted specifically the importance in reading, dissecting, and understanding mathematical definitions as she wrote proofs. Heather stressed how important logical thinking and proof writing was for her future understanding of higher mathematics. Thus, the concept image of concept importance for these participants included indirect abstract algebra concepts, which is vital for professors to consider when teaching the course.

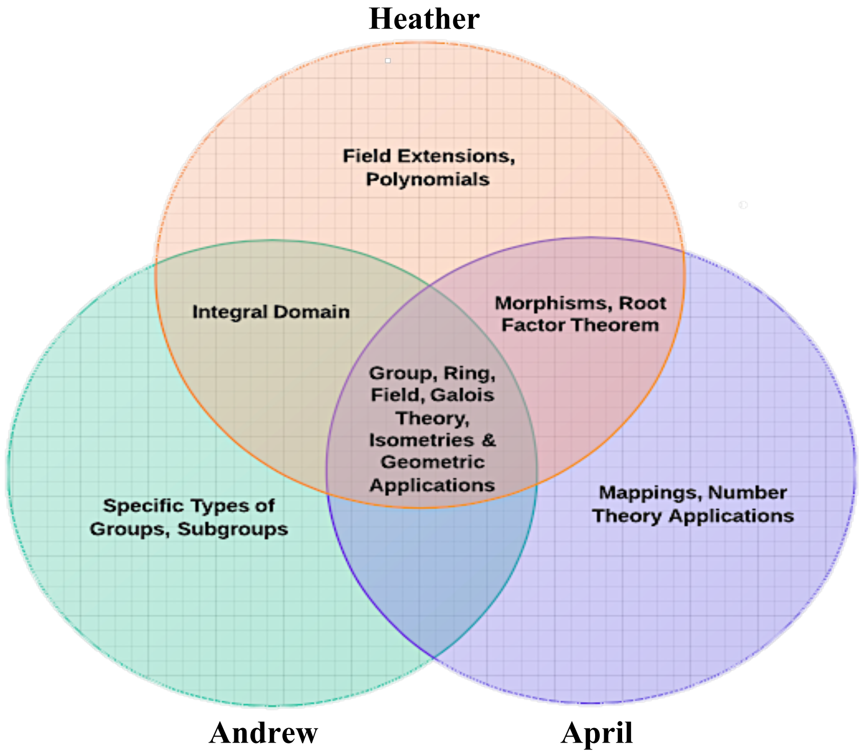


Figure 1: Identified important concepts

In general the participants had difficulty articulating their concept images about content learned in their abstract algebra courses. Despite all of the students acknowledging the intuitive nature of rings, none of them were able to articulate the complete formal concept definition of a ring. April's definition most closely aligned with the formal definition in classifying a ring as a set with two operations following seven axioms, but she could not articulate what were the axioms. Heather's concept image of a ring was similar to April's in that she viewed the algebraic structure in terms of axioms. However, she was unable to articulate a concept definition for a ring due to declared confusion between what those axioms were and how many existed for a ring. Similar to April and Heather, Andrew's concept image of a ring included two operations following certain properties, but his concept image also included the notion of a map, stating: "It is something like you have a map, you have commutativity over addition, associativity over addition, and you have additive identity, you have multiplicative identity."

When asked to describe a group or a field, the students seemed to have less robust concept images of these concepts than a ring. In fact, only April provided a concept definition for a group that closely aligned with the formal concept definition. However, she was unable to provide a concept definition of a field despite being probed several times. Unlike her concept image of a ring and a group that relied heavily on the formal definition, her concept image of a field consisted of types of fields and the concepts taught before and after fields. Heather and Andrew had greater difficulty articulating their concept images and concept definitions. Heather responded, "Gosh. I think I am confused" when asked about a group and "The funny thing is I just totally, I just don't remember what a field was." However, she did eventually attempt to relate her two concept images of a group and a field in stating, "I feel like it (field) has less, no the group has less conditions in order to be a group." This comparison proved to be unhelpful to her understanding of these concepts to which she repeatedly asked me to provide her the formal definition of a group and a field since she could not remember despite earning As in all of her courses. Andrew's concept image of a group relied heavily on his definition of a ring in describing a subset relationship between the two concepts. He stated, "From rings we can get groups. Kind of like subsets of rings are groups because we just have one operation" and "Because rings are the more generated thing with two operations, addition and multiplication, so a group is kind of like throwing one of the operations out." When asked to elaborate more on the definition of a group, Andrew discussed another aspect of his concept image: his perceptions of usefulness of a group in the real world. Ultimately, though, he was unable to provide a concept definition of a group. When asked to define a field he responded, "A field is something I just can't get used to it" and "A field is an integral domain." His concept image of a field also included examples from the number system.

All participants also provided a personal reconstruction concept definition of isometries and geometric applications, whereas no participant could accurately define Galois theory. In discussing the former, Andrew concentrated on functions, April discussed subtopics taught in class, and Heather related isometries to group theory. Andrew was the only participant that elaborated on his concept image of Galois theory, saying, "There are I think orders that tells us any polynomial greater than of degree 2, we cannot necessarily have a formula for factorizing it or something, so that is another I would say that's perhaps an important use of it" Neither April nor Heather offered their concept images of Galois theory despite being asked.

The concept images and concept definitions of perceived relationships between identified important concepts, as seen in the created concept maps, were quite diverse despite the fact that the participants took the same course. Andrew described his concept image of the connections between concepts as a "hierarchical structural" flow chart and a "laying kind of thing" (Figure 2). When asked to describe the arrows drawn between concepts, he admitted to not fully grasping how the concepts in abstract algebra were built upon each other, but he knew they were all somehow related. Surprisingly, Andrew did not draw his subset relationship between a group and a ring that he explained with his concept image of a group. Likewise, he never included applications in his concept map despite concentrating on them during the first half of the interview. Heather, on the other hand, included subset notation in her concept image of concept connections. She described her concept map as a web of concepts with lines denoting concept connections as well as set notation denoting subset relationships (Figure 3). For instance, she used a subset symbol between rings and fields because she understood these concepts as subsets of each other. She contemplated between also including a subset symbol between groups and rings, but she seemed unclear as to how the subset relationship worked in questioned whether a group is a subset of a ring or is a ring a subset of a group.

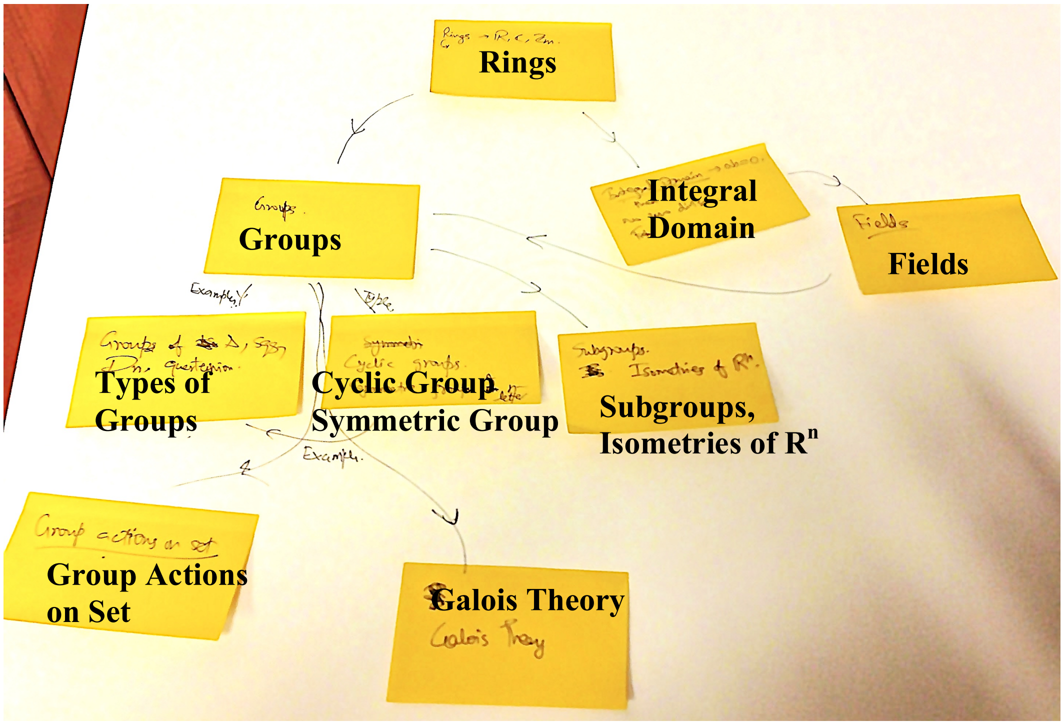


Figure 2: Andrew's Concept Map

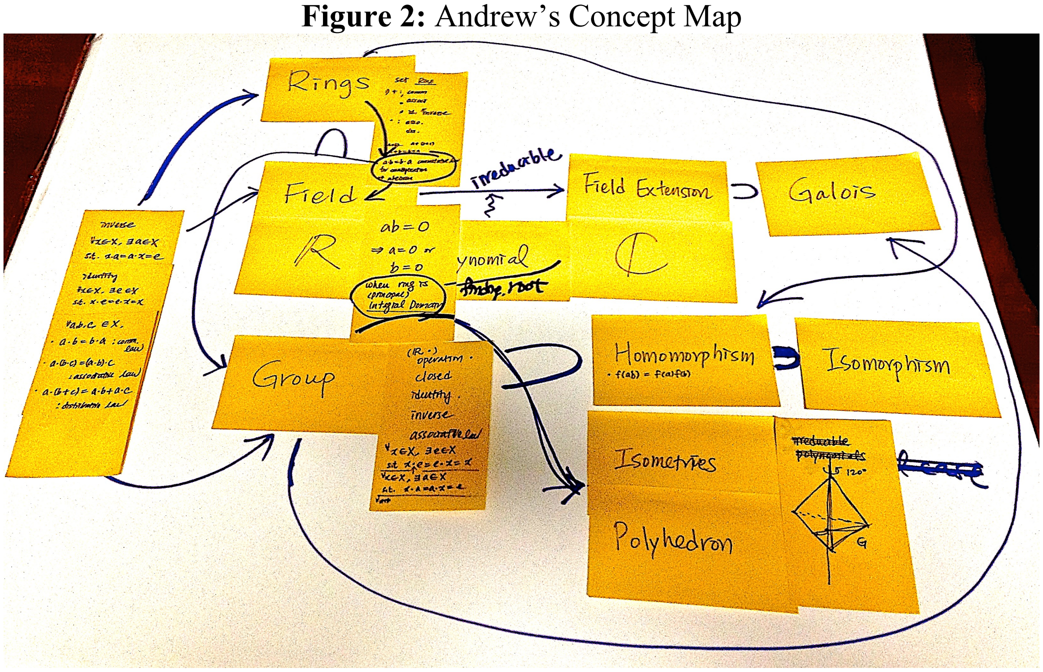


Figure 3: Heather's Concept Map

April's concept image of concept connections was quite different than the other two participants. On her concept map green lines indicated major topics and red lines indicated concept connections. To her, the red arrows were drawn between concepts that had overlapping concept images or similar concept definitions. For instance, when explaining the drawn red arrow between rings and groups she said, "They are related in many ways, especially in the definition when trying to decide whether a set is a ring or is a group. Because they are so closest related to each other." April also described concept connections of major concepts by similar applications of her constructed concept definitions.

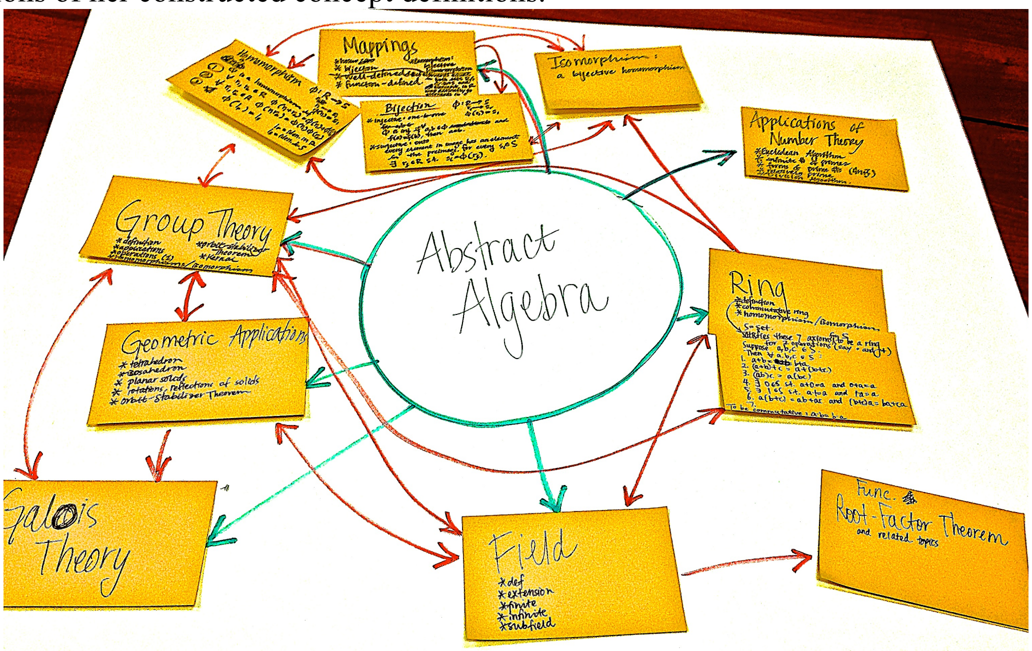


Figure 4: April's Concept Map

Despite the variation in the participants' concept image of concept connections, April and Heather discussed an additional aspect of their concept image: the order the concepts were discussed in class is linked to concept connections. In other words, concepts are related when discussed before and after each other. April explained when asked to elaborate on her concept maps, "The reason I have a bidirectional arrow between fields and rings was because we discussed fields after rings." Likewise, Heather described her arrow between ring and homomorphism, "So that's why I put it together. I just remember using that word ring homomorphism over and over again, so that's why I thought they were connected." This result parallels their concept image of concept importance of these two participants earlier discussed in this paper. Contrary to these results, one student portrayed concept connections in this way: "The main concept of connections is not only based on definitions, but the ways we applied our knowledge of each concept, so for instance, in rings, once we covered the definition of what makes a set a ring, we talked about applications of rings." Consequently, each participant had various reasons for their concept image and concept definition of concept importance and concept connections.

Conclusions

In general, the participants of this study attempted to rely on connections to past mathematical knowledge or real-world applications when articulating their concept images of major abstract algebra topics. In fact, the participants of this study seemed to be searching for any missing connections between identified important concepts and past mathematical knowledge. Heather affirmed:

> Making connections with other courses or ideas, I feel like that it is really hard to do it but it is important and it's helpful. I really wished I knew this before I taught so that I can make better connections in my own teaching. … Because then I would have been able to provide more let's say examples or even provide more opportunities for them to think about things to make connections between the mathematical ideas.

Similarly, April repeatedly discussed connections between abstract algebra and number theory, "I generally like to think about abstract algebra and number theory being two courses that are very closely linked, partly because number theory is more of an application course of abstract algebra and abstract algebra is the more theoretical course of number theory." Likewise, all three participants discussed the applications (or lack thereof) of abstract algebra concepts. Andrew claimed abstract algebra was not useful due to his inability to apply the theorems and definitions to real-life:

> I don't know if it is useful. Like to be honest, I don't know what I am learning. It's like we learn lots of theorems and it's kind of like solving problems but I really don't know if we will ever be able to apply them to real life. Or use them since they are all so… I know that I learned all these concepts, but I don't know if there is any usage out of them and if I am ever going to use them in real life.

Andrew's concept images were typically tied to these applications, so his inability to identify them only hinders his learning. Thus in order for these graduate students to construct robust concept images and accurate concept definitions, there must be connections to past knowledge and real-world applications, which concurs Cook (2012) and Fennema and Franke (1992).

In addition, the difficulties these participants had in articulating their concept images and concept definitions parallels the learning scenarios described in Vinner (1983). April in particular seemed to initially accept the formal concept definition of a ring, group, and field when describing her concept images, but ultimately personally constructed concept definitions of a group and field when her concept images did not match the formal definitions. Heather also seemed to rely on the taught formal concept definitions of a ring, group, and field to construct her concepts images. However, her concept images as well as her concept definitions of a group and a field seemed to be forgotten, which aligns with Vinner (1983) that students whom solely rely on rote memorization to define a concept will eventually forget both. Unlike the two female participants, Andrew attempted to construct his concept images of a ring, group, and field upon previously constructed images of related concepts and the formal concept definitions. In spite of his attempts, he was unable to establish accurate connections between concept images, which caused him difficulty in articulating his concept images and concept definitions. In general, these participants tried to allow the concept definition to form new concept images, but it was not enough to develop accurate images and definitions of abstract algebra concepts.

Implications for Future Research

The results of this research study illustrates for these participants that their constructed concept images and concept definitions differ from the formal definitions taught in class. To put it plainly, students are not constructing the same mathematical meanings behind concepts as expected. Tertiary mathematics professors should then consider how to minimize this mismatch in learning to foster the development of more accurate concept definitions. This follows from Vinner (1983), "revealing the concept images of our students becomes very important for teaching; not only might it give us a better understanding of our students but also it might suggest some improvements to our teaching which formed such wrong concept images." One must provide students with enough examples that form the desired concept image not only in the beginning of the concept image development but throughout the learning process. Elements that cannot constantly being reinforced have a good chance of being forgotten, resulting a distorted concept image. Similarly, this study provides abstract algebra professors a snapshot into what students identify as important to the course and how these concepts are defined. This finding is particularly useful since many mathematics professors may not know what the students are actually learning or not learning in their classes. Furthermore, these identified important concepts may or may not correspond to the professor's identified important concepts. As seen in this study, despite the participants overlapping four major concepts, each constructed very different concept images and concept definitions of these concepts that did not resemble the formal definition. In future work, I hope to utilize the methodology and results of this study to investigate the connections between school algebra and the abstract algebra course. These connections are often missing from classroom instruction even though students would benefit greatly from explicit instruction of them. Likewise, as seen in this study, students desire to establish connections between identified new abstract concepts and past mathematical knowledge.

References

Asiala, M., Brown, A., Kleiman, J., & Mathews, D. (1998). The development of students' understanding of permutations and symmetries. *International Journal of Computers for Mathematical Learning*, *3*, 13-43.

Asiala, M., Dubinsky, E., Mathews, D. M., Morics, S., & Oktaç, A. (1997). Development of students' understanding of cosets, normality, and quotient groups. *Journal of Mathematical Behavior, 16*(3), 241–309.

Brown, A., DeVries, D. J., Dubinsky, E., & Thomas, K. (1997). Learning binary operations, groups and subgroups. *Journal of Mathematical Behavior, 16*(3), 187–239.

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 509-535). Thousand Oaks, CA: Sage.

Cook, J. P. (2012). *A guided reinvention of ring, integral domain, and field*. (Doctoral dissertation). Retrieved from ProQuest. (UMI 3517320)

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics, 27*, 267–305.

Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147–164). New York, NY: Macmillan.

Freedman, H. (1983). A way of teaching abstract algebra. *American Mathematical Monthly, 90*(9), 641–644.

Gallian, J. A. (1990). *Contemporary abstract algebra* (2nd ed.). Lexington, MA: D. C. Heath.

Hart, E. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. In J. Kaput & E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning: Preliminary analyses and results* (pp. 49–63). Washington, DC: Mathematical Association of America.

Hazzan, O. (1999). Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics, 40,* 71-90.

Larsen, S. (2004). *Supporting the guided reinvention of the concepts of group and isomorphism: A developmental research project*. Unpublished Doctoral Dissertation, Arizona State University.

Larsen, S., & Lockwood, E. (2013). A local instructional theory for the guided reinvention of the quotient group concept. *The Journal of Mathematical Behavior*, *32*(4), 726-742.

Leron, U., & Dubinsky, E. (1994). *Learning abstract algebra with ISETL*. New York, NY: Springer-Verlag.

Leron, U., & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly, 102*(3), 227-242.

Leron, U., Hazzan, O., & Zazkis, R. (1995). Learning group isomorphism: A crossroads of many concepts. *Educational Studies in Mathematics, 29*, 153–174.

Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*, (Technical Report IHMC CmapTools 2006-01 Rev 01-2008). Retrieved from Florida Institute for Human and Machine Cognition: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pedersen, J. J. (1972). Sneaking up on a group. *Two Year College Mathematics Journal, 3*(1), 9–

12.
Selden, A., & Selden, J. (1987). Errors and misconceptions in college level theorem proving. In *Proceedings of the 2nd international seminar on misconceptions and educational strategies in science and mathematics*. Ithaca, NY: Cornell University.
Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12,* 151-169.
Taylor, S. J., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings.* New York: Wiley.
Trochim, W. M. K. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning, 12,* 1-16.
Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology, 14*(3), 293-305.
Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*, 101–119.
Zazkis, R., & Hazzan, O. (1999). Interviewing in mathematics education research: Choosing the questions. *Journal of Mathematical Behavior*, *17*(4), 429-439.

PROFESSIONAL DEVELOPMENT AND STUDENT GROWTH ON A STATE MATHEMATICS ASSESSMENT

Melissa L. Troudt
University of Northern Colorado

Rebecca A. Dibbs
University of Northern Colorado

Robert A. Powers
University of Northern Colorado

This is a preliminary report of a study where the purpose was to examine how participation in a mathematics master's program for in-service teachers affected student growth on a state mathematics assessment taking into account student demographic variables. We collected data from school districts for 5 academic years spanning from pre-program through program completion. We conducted a four-step hierarchical multiple linear regression analysis. We can conclude that the combination of teachers' years of participation in the Math TLC, teachers' total years of experience, student ethnicity, and student free and reduced lunch eligibility provided a joint effect on the student growth percentiles on the state mathematics assessment. We explain future plans for continued research on this project.

Key words: in-service teachers, professional development, secondary education, student achievement

It is widely acknowledged that highly qualified teachers are instrumental to student achievement. Therefore, the Mathematics Teacher Leadership Center (Math TLC), an NSF-funded Mathematics and Science Partnership project, has developed and is researching a master's program in mathematics education in response to the call for advanced professional education accessible to in-service teachers. The Math TLC is a blended online and face-to-face, 2-year program that aims to effect change in in-service teachers' content proficiency, cultural competence, and pedagogical expertise for teaching secondary grades mathematics (grades 6 to 12). As such, an important outcome variable for the research team to measure is student achievement. We do not assume that the state mathematics assessment accurately reflects all that was presented to our teachers during the master's program. However, as student performance on state assessments may be factored into teachers' professional evaluations, those that deliver professional development and teachers will be interested in how teacher participation in these professional development programs can play a role in student achievement.

Although measures of teacher preparation and certification are the strongest correlates of student achievement in reading and mathematics (Darling-Hammond, 2000), quantifying what types of professional development have an impact on student achievement still warrants exploration. We seek to explore to what extent teacher participation in the Math TLC master's program affected their students' growth on a state-wide test accounting for student demographic variables. The state tests considered are the Colorado Student Assessment Program (CSAP) and the Transitional Colorado Assessment Program (TCAP)

This is a quantitative causal-comparative study (Gall, Gall, and Borg, 2007) on the state scores of the students of teacher participants of the master's program. We have collected the participant's students' demographic and state assessment data for four years. We also have conducted quantitative observations of the teacher participant's teaching both before and after his enrollment in the program (Hauk, Jackson, & Noblet, 2010; Goss, Powers, & Hauk, 2013) measures of his pedagogical content knowledge for teaching using a written instrument (Hauk, Toney, Jackson, Nair, & Tsay, 2013) throughout his participation in the program, and measures of his intercultural competence pre-program and post-program.

Literature

High quality teachers improve student achievement, but any easily measured credential is at best a barely significant measure of teacher quality (Foster, Toma, & Troske, 2013; Rockoff, 2004), and the pedagogical content knowledge that appears to account for most of the variance in teacher quality is difficult to measure (Dash et al, 2012). Although professional development can help in-service teachers to examine and improve their practice, the conflicting research on the relationship between professional development and student achievement indicates that this relationship may be dependent on the specifics of a given professional development program, which is a challenge for teacher educators given that NCLB does not define what high quality professional development is (Blank & de las Alas, 2009; Huffman, Thomas & Laurenz, 2003; Ross, Hogaboam-Grey, & Bruce, 2006).

Professional development programs that succeed in raising student achievement on standardized tests have focused on specific instructional practice skills or curriculum development and had multiple years of continuous implementation (Hoffman, Thomas, & Laurenz, 2003). The instructional practices that are most likely to lead to student achievement gains following professional development are instruction techniques in a specific content area, like fraction addition or in the use of synchronous formative assessments like stoplight cards in the classroom (McGraner, Van Der Heyden, & Holdheide, 2011). Laura, McMeeking, Orsi, and Cobb (2012) found both significant student achievement gains on the CSAP for teachers that completed a particular professional development program; students' odds of going from not proficient to proficient also increased.

Student achievement scores are a popular way to measure teacher gains, since this is data that can be collected without additional loss of instructional time (Dash et al, 2012; Foster, Toma, and Trotske, 2013; Laura, McMeeking, Orsi, & Cobb 2012). There are several measurement issues inherent in using state standardized testing data, in this case the CSAP/TCAP. Although the test is reliable and the items are psychometrically well constructed (CTB 2010, 2011), the bookmark standard setting process is not tied to any particular curriculum. For researchers, the standard setting can create two levels of disconnect between the professional development and the state assessment; the difference between the program and what is implemented in the classroom and the difference between the classroom and the state exam. Due to this potential validity issue, using state standardized test scores in a linear model will tend to and lower the R-squared value of the model (Karantonis & Sireci, 2006; McGinty, 2006).

However, even when there is significant achievement gains attributable to professional development, student achievement occur after the professional development has ended (Harris & Sass, 2007). One of the difficulties in measuring student achievement gains following professional development is the implementation dip, where student achievement drops because teachers have discarded their old teaching practices but have not mastered the skills in the professional development (Ball, 2004; Busnick & Inos, 1992). Although the implementation dip is part of incorporating professional development into practice, this drop can be minimized by conducting professional development in a single area, following up with teachers in their classroom, and revisiting the same topics in later professional development (Stiles, Loucks-Horslet, Mundry, Hewson & Love, 2009; Zapeda, 2012). The implementation dip lasts up to 18 months after the last professional development session, but recent large scale studies of professional development on middle school and secondary mathematics teachers showed an implementation dip that lasted up to three years before there were measurable student achievement gains (Blank, Smithson, Porter, Nunnaley, & Osthoff, 2006; Busnick & Inos, 1992; Harris & Sass, 2011).

Methods

Setting and Participants.

The Math TLC 2-year master's program is for in-service secondary teachers in mathematics with an emphasis in teaching. About half of the course credits are in mathematics and half are in education. The program delivery blends face-to-face and online formats and is offered jointly between two Rocky Mountain region universities. The primary goals of the program are to develop with in-service teachers a vision of mathematics as a culturally rich subject, increase teachers' pedagogical content knowledge by examination of how students think and learn about mathematics, and expand mathematical content knowledge in topics that extend K-12 mathematics content. So far, 31 teachers from 3 cohorts have successfully completed the program. Our previous research has focused on changes in teachers' pedagogical content knowledge for mathematics (Goss, Powers, Hauk, 2013), and now we are seek to describe student outcomes.

This is a quantitative causal relationship study (Gall, Gall, & Borg, 2007) on the state scores of the students of teacher participants from the first two cohorts to successfully complete the master's program. We requested state assessment and demographic data of the students of 20 teacher participants from 10 districts. We received data from 2 districts pertaining to 9 teacher participants. Five teachers were from the first cohort of teacher participants that completed the program in the summer of 2011, and 4 teacher participants were in the second cohort that completed the program in the summer of 2012. Table 1 summarizes the demographic data of the students of the teachers from the first two cohorts.

	N	*%*		*N*	*%*
Total	2559	100			
Teacher			*Gender*		
TLC1046	361	14.1	Female	1300	50.8
TLC1051	293	11.4	Male	1259	49.2
TLC1081	230	9.0			
TLC1086	207	8.1	*Ethnicity*		
TLC2042	352	13.8	White	1900	74.2
TLC2052	514	20.1	Non-white	659	25.8
TLC2077	337	13.2			
TLC2087	265	10.4	*Free/Reduced Lunch*		
Year in Program			Not Eligible	1688	66.0
0 years (not started)	534	20.9	Eligible	632	24.7
1 year	727	28.4			
2 years	680	26.6			
3 years (completed)	618	24.2			

Table 1: Demographic frequency counts for students of teacher participants.

Data Collection

We collected the student mathematics scores on a state assessment that is administered to students in grades 3-10 in the subjects of mathematics, English/Language Arts, and science in the spring of each academic year. Within each subject, students receive a scale score, a proficiency level, and a growth percentile rating for the exam. Scale scores are conversions from raw scores that represent the same level of achievement regardless of the year in which the test was administered. The growth percentile is a representation of each student's progress comparing each student's current achievement to students in the same grade throughout the state who scored similarly in past years (Bettebenner, 2009). The growth percentile is only reported if the student completed the assessment in two consecutive years. The first cohort of teacher participants began the Math TLC in the summer of 2009, and the second cohort completed the program in the summer of 2012. Therefore, we collected the scores of state assessments that were administered in the spring of 2009 through the spring of 2013. The state is in the process of implementing new academic standards. The state test used through 2011, CSAP, assessed the previous standards. The test used from 2011 through 2014, TCAP, was designed to measure standards common between the new and old standards. In practice, districts continue to compare across years and growth percentiles calculated for student performance on the TCAP took into account their performance on the CSAP in prior years.

The writers of the state-administered instrument discuss their efforts to establish content validity by having content-area specialists, teachers, and assessment experts develop a pool of items that evaluated the state's assessment framework in each grade and content area. Measures of reliability including internal consistency and interrater reliability are calculated each year after the assessments are administered to students across the state. The scores of the students of our teachers are included in this calculation. The state reports the mathematics test showed good internal consistency; Cronbach's alphas range from 0.92 to 0.94. They report interrater reliability kappas range from 0.66 to 0.94 (Colorado Department of Education, 2009, 2010, 2011, 2012).

We followed school district protocols to obtain student demographic and assessment data. We requested students' demographic data, state scale mathematics scores, state mathematics proficiency levels, student growth percentiles in mathematics, and state English and Language Arts scale scores. For each teacher, for each of the given years, we requested all data for students in that year plus all data for those same students for the previous year. The student demographic data on which we chose to focus was student gender, student ethnicity, and student free and reduced lunch eligibility. We chose to focus on these three variables because they were easily obtained and they could account for variability in student achievement. We also collected teachers' total years of teaching experience. We included years of experience as an independent variable as more experience in teaching may be correlated to higher student achievement. We wish to determine if teacher participation in and completion of the professional development affects student growth beyond what can be explained by teachers' total years of experience.

We obtained students' mathematics scale scores, mathematics growth percentiles, and demographic data pertaining to each given year, but we did not obtain the data linking students' current year scores to previous year scores. We considered the state assessment scores of the teachers' students in grades 6-10 from five academic years: (0) prior to the teachers beginning the program, (1) the teachers' first year in the program, (2) the teachers' second year in the program, and (3) the year following the teachers' completion of the program, and (4) two years following the teachers' completion of the program. At this point, we only have data on all five levels for three teacher participants; we have data on four levels for all teachers.

Data Analysis

Because we did not obtain linked data to students' previous years' scores, we chose to focus on students' mathematics growth percentiles as the outcome variables. The purpose of this analysis was to determine if teacher participation in the Math TLC can explain variance in student growth on the state assessment accounting for student demographic variables and teachers' years of experience. For variance explanation, Pedhazur (1997) recommended using hierarchical linear analysis.

Predictor variables included student gender, student ethnicity, student free and reduced lunch eligibility, the teachers' total years of teaching experience, and the year of enrollment in the Math TLC of the teacher at the time of the student assessment. Gender is a dichotomous variable; to prevent having very small groups, we entered ethnicity (white, non-white) and free and reduced lunch eligibility (free and reduced lunch eligible and non-eligible) as dichotomous variables. We effect coded the non-continuous, predictor variables. Multiple regression analyses including categorical variables should account for possible interactions among independent variables. It may be that the joint effect of ethnicity and gender explains more variance than ethnicity alone; Pedhazur (1997) calls this a joint or multiplicative effect. For this reason, we calculated interaction terms among all categorical variables.

We performed hierarchical linear regression. Step one included gender, ethnicity, free and reduced lunch eligibility, and years of teaching experience. Step two included the primary predictor variable of interest, the four vectors from the effect coding of the five levels of year in the Math TLC. Higher level steps included interaction terms (products of two variables were entered at step three, three variables at step four, etc.). Testing assumptions of linear regression, we found the regression showed a non-linear relationship; therefore, we transformed the dependent variable by taking the arcsine of the square root of the student mathematics growth percentile (Sheskin, 2003). We performed the hierarchical analysis with the steps above and performed all-subsets tests to determine if the interaction terms held a significant effect (Pedhazur, 1997). We eliminated non-significant product terms from the model and ran hierarchical analysis on the smaller model.

Results

We first report descriptive statistics showing student mean growth percentiles on the state mathematics assessments for each year the teachers were enrolled in the program in Table 2. We see that mean growth percentile increased slightly while teachers were enrolled in the first year of the program, decreased slightly, and means increased again only two years following teacher completion of the program. No changes appear to be significant.

Year in Program	*N*	*M*	*SD*
0 (pre-program)	398	48.0	28.76
1 (mid-program)	550	52.0	29.23
2 (mid-program)	615	48.8	29.20
3 (1 year post-program)	579	48.5	30.42
4 (2 years post-program)	116	52.2	30.73

Table 2. Student growth percentile means by year of teacher enrollment in the Math TLC.

In the all-subsets test of variables, only the joint effects of free and reduced lunch eligibility with year in the program; ethnicity with years of teacher experience; ethnicity with year in the program; free and reduced lunch eligibility with year in program; the product of ethnicity, years of experience, and year in the program, and the product of free and reduced lunch eligibility, years of experience and year in the program. All variables considered in the model are given in Table 2. Table 3 summarizes the correlations among the main effect predictor variables; joint effect variables are not included. Note that Y1 indicates the first year the teacher was enrolled in the program, Y2 the second year, etc. Because categorical variables were effect coded and groups were not of equal size, correlations among groups are non-zero.

Predictor Variables	
1) Ethnicity	14) Ethnicity-Y1
2) Gender	15) Ethnicity-Y2
3) Free/Reduced Lunch Eligibility	16) Ethnicity-Y3
4) Years of teaching	17) Ethnicity-Y4
5) Year1	18) F/RLunch*YrsTeaching*Y1
6) Year2	19) F/RLunch*YrsTeaching*Y2
7) Year3	20) F/RLunch*YrsTeaching*Y3
8) Year4	21) F/RLunch*YrsTeaching*Y4
9) F/RLunch*Y1	22) F/RLunch*YrsTeaching*Ethnicity*Y1
10) F/RLunch*Y2	23) F/RLunch*YrsTeaching*Ethnicity*Y2
11) F/RLunch*Y3	24) F/RLunch*YrsTeaching*Ethnicity*Y3
12) F/RLunch*Y4	25) F/RLunch*YrsTeaching*Ethnicity*Y4
13) Ethnicity-YrsTeaching	

Table 2. Predictor variables entered into the hierarchical linear analysis.

	1	2	3	4	5	6	7	8
1) Ethnicity	--	-.01	.39	-.01	-.02	-.03	.00	.06
2) Gender		--	.00	-.04	.02	-.02	.00	.02
3) Free/Reduced Lunch Eligibility			--	-.01	.03	.01	.05	.13
4) Years of Teaching				--	.01	.10	.17	.24
5) Year1					--	.40	.41	.62
6) Year2						--	.39	.62
7) Year3							--	.62
8) Year4								--

Table 3. Correlations between main effect predictor variables.

Due to the transformation of the dependent variable and the effect coding of the categorical independent variables, the partial slopes in the regression model should not be interpreted in the usual sense (Pedhazur, 1997). Additionally, since each main effect variable apart from gender was a factor in some significant joint effect, we cannot interpret the main effects. Instead, we consider if inclusion of variables at each level explained significantly more variance in the dependent variable than the previous levels.

The hierarchical multiple regression revealed that at Stage one, Free and Reduced Lunch Eligibility contributed significantly to the regression model, and the block of predictor variables ethnicity, gender, free/reduced lunch eligibility, and teacher's years of experience accounted for 2% of the variance in the transformation of student growth percentile (*F-change*(4,2253) = 10.78, $p < .001$). Adding the block of variables representing teachers' years of participation in the Math TLC explained only an additional 0.4% of the variance, but the F Change statistic was found to be significant, *F-change* (4,2249) = 2.42, $p = .047$. Adding the joint effects of free/reduced lunch eligibility with the years of the program, ethnicity with total years of teaching experience, and ethnicity with the year in the program explained an additional 1% of the variance (*F-change* (9,2240) = 2.37, $p = .005$). Finally adding the joint effects of ethnicity with years of experience with year in the program and free and reduced lunch eligibility with years of experience with ethnicity with year of the program explained an additional 1% of the variance (*F-change* (8, 2232) = 2.57, $p = .009$). When all the predictor variables were included in the final stage of the model, all variables accounted for 4.2% of the variance in student growth percentile.

Variable	Step 1	Step 2	Step 3	Step 4
Step 1				
Ethnicity	-.02	-.02	-.03	.12
Gender	-.02	-.02	-.13	-.03
Free-Reduced Lunch Eligibility	-.12***	-.13***	-.01***	-.13***
Teacher's Years of Experience	.01	.01	-.01	-.01
Step 2				
Year1		.03	-.04	-.02
Year2		-.05	.01	-.04
Year3		.00	.07	.02
Year4		.06	-.09	.06
Step 3				
F/RLunch*Y1			-.09**	.06
F/RLunch*Y2			.05	.07
F/RLunch*Y3			-.03	.15
F/RLunch*Y4			-.02	-.22*
Ethnicity*YrsTeaching			-.10	-.14*
Ethnicity*Y1			.03	-.16*
Ethnicity*Y2			.00	-.12
Ethnicity*Y3			.06	.14
Ethnicity*Y4			-.01	.18
Step 4				
F/RLunch*YrsTeaching*Y1				-.16
F/RLunch*YrsTeaching*Y2				-.01
F/RLunch*YrsTeaching*Y3				-.17*
F/RLunch*YrsTeaching*Y4				.18*
F/RLunch*YrsTeaching*Ethnicity*Y1				.19*
F/RLunch*YrsTeaching*Ethnicity*Y2				.11
F/RLunch*YrsTeaching*Ethnicity*Y3				-.09
F/RLunch*YrsTeaching*Ethnicity*Y4				-.17
R	.14	.15	.18	.21
R^2	.02	.02	.03	.04
R^2-*change*	.02	.00	.01	.01

$N = 2134$, $*p < .05$, $**p < .01$, $***p < .001$

Table 4. Beta values and R^2 values from hierarchical regression

Discussion and Conclusions

We sought to describe how participation in the Math TLC master's program could affect student achievement on a state mathematics assessment taking into account other demographic variables. When interaction terms of a multiple regression are statistically significant, it is not meaningful to do multiple comparisons among main effects. Instead we considered the changes in the R-squared values. We found that years of participation in the program were significant contributors to student growth percentile as part of joint effects that also included free and reduced lunch eligibility, total years of teaching experience, and ethnicity. Prior to the inclusion of these joint effects, only students' free and reduced lunch eligibility contributed significantly to the model.

In Table 2, we see that the means of student mathematics growth percentiles for the second year teachers were enrolled in the program and the first year following their program completion were lower than they were pre-program. They began to increase two years after the teachers completed the program. This implementation dip is consistent with the findings from past studies (Harris & Sass, 2007; Ball, 2004; Busnick & Inos, 1992).

Even after the inclusion of all predictor variables, only four percent of the variance in student growth percentile on the state mathematics exam was explained. The low R-squared value may be partially explained by the validity issue of the state assessment not aligning with any one curriculum (Karantonis & Sireci, 2006; McGinty, 2006). Additionally, we interpret that student growth percentile is a variable that cannot be adequately predicted by the given student demographic data and teacher participation in the Math TLC alone. Future research efforts will include other teacher variables including years of other professional development, observation data of teachers, and data from measures of the teachers' pedagogical content knowledge.

We can conclude that the combination of teachers' years of participation in the Math TLC, teachers' total years of experience, student ethnicity, and student free and reduced lunch eligibility provided a joint effect on the student growth percentiles on the state mathematics assessment. Further investigations are required to interpret these joint effects including how well the variable of teachers' total years of experience predicts student growth across these groups.

Moreover, we will continue to collect student data. Collecting data from the 2014 version of the state assessment will provide data pertaining to 3 years following the teachers from the first cohort completion of the program, 2 years following the second cohort's completion of the program, and 1 year following a third cohort's completion of the program. Continuing to collect this data will allow us to assess the program's longer-term effects on student outcomes and potential differences in effects by cohort.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. DUE0832026. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Ball, D. L. (2004). Facilitation of Action Learning Groups: An Action Research Investigation (Doctoral dissertation, Deakin University).

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Educational Measurement: Issues and Practice, 28*(4), 42-51.

Blank, R. K., & de las Alas, N. (2009). Effects of teacher professional development on gains in student achievement. *Peabody Journal of Education*, *77*(4), 59-85.

Blank, R. K., Smithson, J., Porter, A., Nunnaley, D., & Osthoff, E. (2006). Improving instruction through schoolwide professional development: Effects of the data-on-enacted-curriculum model. *ERS spectrum*, *24*(2), 9-23.

Busick, K. U., & Inos, R. H. (1992). Synthesis of the Research on Educational Change: Implementation Phase.Honolulu, HI: Pacific Region Education Laboratory.

Colorado Department of Education (2009, 2010, 2011, 2012). Colorado State Assessment Program Technical Report. Available at http://cde.co.us.

Colorado Department of Education (2012). Transitional Colorado Assessment Program Technical Report. Available at http://cde.co.us.

CTB McGraw-Hill (2010, October). *Colorado student assessment program: Technical report 2007*. Monterey, CA: McGraw-Hill. Retrieved from http://www.cde.state.co.us/cdeassess/publications.html

CTB McGraw-Hill (2011, October). *Colorado student assessment program: Technical report 2011*. Monterey, CA: McGraw-Hill. Retrieved from http://www.cde.state.co.us/cdeassess/publications.html

Dash, S., de Kramer, R. M., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of Online Professional Development on Teacher Quality and Student Achievement in Fifth Grade Mathematics. *Journal of Research on Technology in Education*, *45*(1), 1-26.

Foster, J. M., Toma, E. F., & Troske, S. P. (2013). Does Teacher Professional Development Improve Math and Science Outcomes and Is It Cost Effective?. *Journal of Education Finance*, *38*(3), 255-275.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007) Educational research: An introduction. Boston: Pearson Education.

Goss, M., Powers, R., & Hauk, S. (2013). Identifying change in secondary mathematics teachers' pedagogical content knowledge. In S. Brown, G. Karakok, K. H. Roh, & M. Oehrtman, *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 248-257). Denver, CO: Online at http://sigmaa.maa.org/rume/Site/Proceedings.html.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, *95*(7), 798-812.

Hauk, S., Jackson, B., & Noblet, K. (2010). No teacher left behind: Assessment of secondary mathematics teachers' pedagogical content knowledge. *Proceedings for the 13th Conference on Research in Undergraduate Mathematics Education (electronic)*. PDF available at sigmaa.maa.org/rume/crume2010/Archive/HaukNTLB2010_LONG.pdf

Hauk, S., Toney, A., Jackson, B., & Nair, R., Tsay, J.-J. (2013). Illustrating a theory of pedagogical content knowledge for secondary and post-secondary mathematics instruction. In S. Brown, G. Karakok, K. H. Roh, & M. Oehrtman, *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 308-322). Denver, CO: Online at http://sigmaa.maa.org/rume/Site/Proceedings.html.

Huffman, D., Thomas, K., & Lawrenz, F. (2003). Relationship between professional development, teachers' instructional practices, and the achievement of students in science and mathematics. *School Science and Mathematics*, *103*(8), 378-387.

Karantonis, A., & Sireci, S. G. (2006). The Bookmark Standard-Setting Method: A Literature Review. *Educational Measurement: Issues and Practice*, *25*(1), 4-12.

Laura, B., McMeeking, S., Orsi, R., & Cobb, R. B. (2012). Effects of a teacher professional development program on the mathematics achievement of middle school students. *Journal for research in mathematics education*, *43*(2), 159-181.

Lewis, D. M., & Haug, C. A. (2005). Aligning policy and methodology to achieve consistent across-grade performance standards. *Applied Measurement in Education*, *18*(1), 11-34.

McGraner, K.L., Van Der Heyden, A. & Holdheide, L. (2011, January). Preparation of Effective Teacher in Mathematics: A TQ Connection Issue Paper on Applying the Innovative Configuration to Mathematics Teachers Preparation. Retrieved July 8, 2013, from http://www.tqsource.org/pdfs/TQ_IssuePaper_M...

McGinty, D. (2005). Illuminating the" Black Box" of Standard Setting: An Exploratory Qualitative Study. *Applied Measurement in Education*, *18*(3), 269-287.

Pedhazur, E.J. (1997). *Multiple Regression in Behavioral Research: Explanation and Prediction,Third Edition.* Toronto: Thomson Learning, Inc.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247-252.

Ross, J. A., Hogaboam-Gray, A., & Bruce, C. (2006). The impact of a professional development program on student achievement in grade 6 mathematics. *Journal of mathematics teacher education*, *9*(6), 1-27.

Sheskin, D. (2003). *Parametric and Nonparametric Statistical Procedures. Third Edition.* Chapman & Hall/CRC.

Stiles, K. E., Loucks-Horsley, S., Mundry, S., Hewson, P. W., & Love, N. (Eds.). (2009). *Designing professional development for teachers of science and mathematics*. Sage Publications.

Zepeda, S. J. (2012). *Professional development: What works*. Eye on Education.

SLOPE AND DERIVATIVE: CALCULUS STUDENTS' UNDERSTANDING OF RATES OF CHANGE

Jennifer G. Tyne
University of Maine

Studies have shown that students have difficulties with the concepts of slope and derivative, especially in the case of real-life contexts. I used a written survey to collect data from 74 differential calculus students. Students answered questions about linear and nonlinear relationships and interpretations of slope and derivative. My analysis focused on students' understanding of slope as a constant rate of change and derivative as an instantaneous rate of change, and what these meant in the context of the problems. Preliminary results indicate that students have more success with slope questions than derivative questions (McNemar's test, $p<0.05$), and that while students correctly use the slope of a linear relationship to make predictions, they do not demonstrate an understanding of the derivative as an instantaneous rate of change and an estimate of the marginal change.

Key words: Calculus, Derivative, Rates of Change, Slope, Student Understanding

Introduction and Research Questions

America's international competitiveness in the areas of science and mathematics is undermined by the declining mathematics and science literacy of Americans (Seymour & Hewitt, 1998). We need to improve mathematics and science education, with one goal being to produce more science, technology, engineering, and mathematics (STEM) majors (Holdren & Lander, 2012). Combinations of task forces, conferences, commissions, and workgroups, all sponsored by a variety of different organizations, have focused on the causes and consequences of low interest in, and high attrition from, mathematics and science. One such focus is on the pedagogical context of undergraduate learning, and the unmet needs of students (Seymour & Hewitt, 1998). The focus of this study is on one slice of this issue, namely calculus students' understanding of some key concepts needed to succeed in calculus and higher-level mathematics.

Examining calculus students' understanding of slope and derivative as rates of change in the context of real life situations is important. Both slope and derivative are essential to understanding central themes in mathematics, but the "rate of change" concept is not well understood by students (Bezuidenhout, 1998). Understanding slope as a rate of change is the foundation that calculus students need to bring to the learning of derivatives as instantaneous rates of change (Hackworth, 1994).

It is important for the mathematics community to understand students' understanding of slope coming into calculus, and expand on that knowledge in teaching the derivative. "If students do not understand average rate of change, it is hard to imagine they have anything but a superficial understanding of instantaneous rate of change" (Hackworth, 1994, p. 154). And not only must students understand instantaneous rates of change, they must have an understanding of the continuously changing rates, as well as strong covariational reasoning skills to interpreting dynamic situations surrounding functions (Carlson et al., 2002).

My research questions surround the interpretation and use of slope and derivative in the context of real life situations:

- Is there a relationship between calculus students' understanding of slope and their understanding of derivative? Specifically, do students' abilities to interpret the slope as a constant rate of change make them more likely to be able to interpret the derivative as an instantaneous rate of change?

- Do students correctly use the slope and derivative to make valid predictions from models?

By "interpret" I mean to provide a description of the meaning in the context of the problem. And by "use the slope and derivative" I mean do students understand the difference between a constant rate of change (which can be used to interpret change at any x-value) and an instantaneous rate of change (which is only valid for a specific x-value, and can only be used to make predictions around that x-value). Lastly, by "real life situations" I mean application problems that model realistic circumstances. Such applications require students to be able to translate from the context to the abstract level of calculus and then back to the context, skills that require conceptual knowledge (White & Mitchelmore, 1996). "Not only do real-world situations provide meaningful opportunities for students to develop their understanding of mathematics, they also provide opportunities for students to communicate their understanding of mathematics" (Stump, 2001, p. 88).

My researchable questions are directly related to my larger question. While the bigger issue surrounds student understanding of the large concepts of slope and derivative, I focus my research on the understanding and interpretation of each as a rate of change in the context of a real life situation. My focus is on linear and non-linear, one-variable relationships, concepts that are reasonable for first-year calculus students. The study builds off research around student understand of slope and rate of change (Barr, 1980; Barr 1981; Orton 1984; Stump, 2001), student understanding of derivatives (Bezuidenhout, 1998; Ferrini-Mundy & Graham, 1994; Zandieh, 2000), student understanding of the rate of change of linear and non-linear functions (Orton, 1983), and how student knowledge of rates of change affect their conceptual knowledge of the derivative (Hackworth, 1994).

We know that far too many students start in STEM majors, only to drop out due to experiences in early courses (Holdren & Lander, 2012). We must better understand students' knowledge coming into calculus, and how that knowledge can adversely affect their success in calculus, in order to provided a successful calculus experience (and a higher chance they will continue in a STEM field).

Student Understanding of Slope, Derivative, and Rates of Change

Theoretical Perspective

In order to research student understanding of slope and derivative, data were collected through a written instrument. The focus is on a detailed analysis of student understanding of a few key concepts, gained from direct student responses. This approach, which is consistent with a cognitive, theoretical perspective, is well established in the mathematics education community (Siegler, 2003).

This cognitive approach requires certain assumptions, for example the assumption that students are making sense of the tasks in front of them based on their experiences, and that their answers are rational and subject to explanation (Ferrini-Mundy & Graham, 1994). My goal is to understand how individual students are thinking about the ideas of slope and derivative, so I built off the current research by utilizing this theoretical perspective.

Student Understanding of Slope

Research has documented difficulties students have with the concept of slope (Barr, 1980; Barr, 1981; Stump, 2001). Stump (2001) found that while high school students tended to understand slope in functional situations, "many students had trouble interpreting slope as a measure of rate of change" (p. 81). Stump (2001) also found that students tend to think of slope as an angle instead of a ratio and recommended that teachers focus instruction on including a "deliberate mechanism for helping students link the notion of angle to the notion of ratio" (p. 87).

The use of real-world examples to assist in the understanding of slope is wildly used by teachers, both in representing slope in a physical situation (such as a ski slope or wheelchair ramp) and slope in a functional situation (such as population versus time) (Stump, 2001). In the latter functional situation, slope is a measure of rate of change, requiring proportional reasoning, another well-documented difficult concept for students (Orton, 1984).

Student Understanding of Derivative

We want students to not only be able to calculate derivatives, but to understand their meaning in the context of real life situations. Many students come to calculus with a very primitive understanding of functions (Ferrini-Munday & Graham, 1994). And while studies have shown that students' procedural skills for calculating derivatives are often adequate, their conceptual and intuitive understanding of the derivative is lacking (Orton, 1983).

Monk looked at students' understanding of functions from two approaches – point wise and across time. Point wise understanding is what students first attain in their learning about functions, thinking of particular values of the independent variable corresponding to particular values of the dependent variable. But, in calculus, students must have "across-time" understanding of functions, where changes in one variable lead to changes in another variable (Monk, 1994). Researchers have found that students' lack the understanding necessary to deal with these co-varying quantities efficiently, thus not grasping the across time understanding (Bezuidenhout, 1998). Many of the questions on the instrument used in the present study require the across-time understanding of functions.

The concept of the derivative can be represented graphically as the slope of a tangent line, verbally as the instantaneous rate of change, physically as velocity, and symbolically as the limit of the difference quotient (Zandieh, 2000). Much research has been done on the graphical understanding of the derivative (Asiala et al., 1997; Berry & Nyman 2003; Ubuz, 2004), but very little about the verbal interpretation of the derivative as a rate of change. Verbal interpretation was a focus in the present study.

Student Understanding of Rates of Change

Rates of change are the overarching connection between the concepts of slope and derivatives. While much attention is focused on the specific concept of slope in algebra class, the more general concept of rate of change is often not emphasized, and is not well understood by students (Orton, 1984). Hackworth (1994) focused on calculus students' understanding of rate of change, and how their understandings were affected by the instruction of the derivative. She found that the instruction failed to substantially change students' reasoning about rate situations, and she found that the students who did poorly in calculus seem not to understand rate of change deeply. Hackworth also saw that students entering calculus had a weak understanding of rates of change, and that their understanding was relatively unchanged after derivative instruction, and in fact that regardless of the content of the course "what students assimilated was largely irrelevant to their understanding of rates of change" (p. 159).

Prior research found that students are most successful with rate problems involving time, but that more focus should be on helping students form connections among rates involving time, and rates not involving time (Stump, 2001). Students demonstrate the ability to calculate slope but without real understanding of how the calculations relate to the more general concept of rate of change (Barr, 1981). Such rote learning could likely extend into calculus, where the student might be able to learn the techniques, but will be unlikely to understand the concepts (Barr, 1980). Even in studying student understanding of the Fundamental Theorem of Calculus, the difficulties in understanding were often tied to poorly developed understanding of rates of change (Thompson, 1994).

In the graphical understanding of a rate of change, students must recognize the difference between the rate of change of a straight line and a curve, but this was not evident in Orton's study (Orton, 1983). Orton found through interviews that many calculus students do not think about rate of change anymore, losing the connections to the understanding as they moved on to higher level mathematics courses (Orton, 1983).

My Research Motivation

By building off the current research, this study is designed to focus on student understanding of slope, derivative, and rates of change. It also stems from my experiences in the classroom, in both a general education algebra course and a first semester calculus class. In the algebra course, developed to provide students with an alternative to a traditional algebra class, we ask students to do more than just use linear relationships. We want them to understand the slope and what it means in the context of the problem. For example, given an equation such as C = 9.8g + 750, where C is the cost in dollars to produce g gallons of a chemical, students are asked for the slope, the units on the slope, and the interpretation of what the slope means in the context of the problem (in this case, as we increase the number of gallons produced by 1 gallon, the cost increases by $9.80) (Franzosa & Tyne, 2010).

Similarly, we often ask calculus students to interpret the derivative in similar ways. For example, given the C = f(g) is the cost in dollars of producing g gallons of the chemical, what are the units on f'(g)? And, what does f '(200) = 6 represent? In this case, when the number of gallons produced is 200, the cost is increasing at a rate of $6 per gallon (Hughes-Hallet, 2013). But students can answer these types of questions with just a point-wise view of functions. Do they understand the derivative as a function that changes?

We currently know from research that students have a difficult time with understanding slope as a rate of change, with across-time understanding of functions, and with understanding the derivative as an instantaneous rate of change. What there has not been much research on, however, is students' verbal interpretation of the derivative and slope as a rate of change, and students' understanding of the differences in making predictions involving constant rate of change and instantaneous rate of change.

Research Design

Setting

The setting was differential calculus (first semester calculus) classes at a pubic university in the Northeast. The participants were students in two sections in fall 2013; 84 students participated from two sections that met three times per week in a lecture setting with a faculty instructor and twice per week in recitation with a graduate student teaching assistant. Of the 84 students, 74 completed the survey fully and are included in this study. The two sections had different faculty instructors and graduate teaching assistants. Students completed the surveys during class time, approximately 80% through the first semester calculus course. There is only one type of calculus course taught at this university and thus the course enrolls a mixture of engineering students, non-engineering mathematics and science students, and a few students from other disciplines where majors require differential calculus. Over 50% of the students have seen calculus in high school, and all needed to either pass the mathematics placement exam or successfully complete precalculus at the University with a C or better to gain enrollment into differential calculus.

Data Collection

The survey instrument consisted of questions about slope and derivatives (see Figure 1), including questions about linear and nonlinear relationships between the dosage of a drug as a function of the weight of a patient. Neither variable represents time, and therefore the interpretation is slightly different for students than questions where the independent variable is time. Prior research has shown that students are most successful with rate problems involve time, the most intuitive type of rate (Stump, 2001).

The questions are not mechanical in nature and therefore do not assess computational skills; instead, they are questions about students' interpretation of slope and derivative, and therefore try to uncover their conceptual knowledge about these topics.

For certain drugs, the amount of dose given to a patient, D (in milligrams), depends on the weight of the patient, w (in pounds).

1. Assume that D(w) is a linear function with a slope equal to 2 (m = 2).
   a. What are the units on the slope, m = 2?
   b. Explain what this slope (m = 2) means in the context of the problem.
   c. Using the slope (m = 2), Jodi predicts that a patient's dose will increase by 2 mg when the patient's weight changes from 140 pounds to 141 pounds. Do you agree with her reasoning? Explain.
   d. Based on the linear model, a nurse accurately gave a patient a dose of 300 mg. Her next patient is twenty pounds heavier and she reasons that she must increase the dose by 40 mg (2 mg for each pound of weight) for a total dose of 340 mg. Do you agree with her reasoning? Explain.
2. Now, assume D(w) is a non-linear function.
   a. What are the units on $\frac{dD}{dw}$? (also known as $D'(w)$)
   b. Explain the meaning of the statement $D'(140) = 2$ in the context of the problem.
   c. Using the fact that $D'(140) = 2$, Jodi predicts that a patient's dose will increase by 2 mg when the patient's weight changes from 140 pounds to 141 pounds. Do you agree with her reasoning? Explain.
   d. A nurse accurately gave a 140-pound patient a dose of 300 mg. Her next patient is 160-pounds and she reasons that since $D'(140) = 2$, she must increase the dose by 40 mg (2 mg for each pound of weight) for a total dose of 340 mg. Do you agree with her reasoning? Explain.

Figure 1. Survey Instrument

In textbooks and in instruction, when focus is given to students' understanding of slope and derivative, usually the questions asked are similar to 1a, 1b, 2a, and 2b (Figure 1). These questions address units (Bezuidenhout, 1998) and point-wise interpretation of rates of change (Monk, 1994). However, Monk (1994) argues that "across-time understanding of functions is critical to an understanding of calculus" (p. 9) and so other questions were included in the survey.

Because of the need to gather data on across-time understanding, the survey included questions 1c, 1d, 2c, and 2d. Analysis of data from these questions is a focus of the present study. The linear questions (1c and 1d) are posted to gain an understanding of students' knowledge of predictions based on linear change. An understanding of the linear change as a constant is necessary to successfully answer these questions. The questions about nonlinear relationships (2c and 2d) are more complex. In order to answer these questions, students must understand the derivative as an estimate of marginal change, and that the derivative is an instantaneous rate of change that cannot be used to make predictions at other input values.

For each of the questions used as data sources for the present study (1c, 1d, 2c, and 2d), I have included what a "ideal knower" would answer, what students would be thinking about while solving the task, and what the question is designed to give information about. These descriptions were used to inform the data analysis (described later).

#1: Assume that D(w) is a linear function with a slope equal to 2 (m = 2).

c. Using the slope (m = 2), Jodi predicts that a patient's dose will increase by 2 mg when the patient's weight changes from 140 pounds to 141 pounds. Do you agree with her reasoning? Explain.

The ideal knower would respond, "yes," by understanding that a slope of 2 represents the increase in milligrams per pound, and that it is a constant rate of change. As the pounds increase by 1, the dosage increases by 2 mg. This question begins to get at students' across-time understanding of functions, as they have to understanding how the dependent variable changes as the independent variable increases by one.

d. Based on the linear model, a nurse accurately gave a patient a dose of 300 mg. Her next patient is twenty pounds heavier and she reasons that she must increase the dose by 40 mg (2 mg for each pound of weight) for a total dose of 340 mg. Do you agree with her reasoning? Explain.

The ideal knower would respond, "Yes, I do agree," and explain that the increase of 2 milligrams per pound is constant and would be applied to the twenty-pound increase. This question is designed to get at students' knowledge of the slope as a constant rate of change, and how it can therefore be applied to any input values.

#2: Now, assume D(w) is a non-linear function .

c. Using the fact that $D'(140) = 2$ *, Jodi predicts that a patient's dose will increase by 2 mg when the patient's weight changes from 140 pounds to 141 pounds. Do you agree with her reasoning? Explain.*

The ideal knower could respond one of two ways. First, it would be appropriate to respond "Yes, I do agree," by understanding that the instantaneous rate of change can be used as a prediction for the marginal change, or for input values very close to the input value of the derivative. Or, another appropriate response would be "No, I don't agree," by showing an understanding that the function is non-linear so we do not know how different the actual value would be from the tangent line prediction. This problem is designed to get information about students' understanding of the instantaneous rate of change's use in predicting marginal change. The important thing for students to show an understanding about is that the non-linear nature of the function means the derivative gives an estimate of the change (and because information is not given about the type of non-linear function, we are not sure how much error is involved).

d. A nurse accurately gave a 140-pound patient a dose of 300 mg. Her next patient is 160-pounds and she reasons that since $D'(140) = 2$ *, she must increase the dose by 40 mg (2 mg for each pound of weight) for a total dose of 340 mg. Do you agree with her reasoning? Explain.*

The ideal knower would respond, "It is not a valid prediction because 2 milligrams per pound is the instantaneous rate of change for a 140-pound person. Because the function is non-linear, one can not use the instantaneous rate of change to make a prediction so far away from 140-pounds." This ideal knower would understand that the instantaneous rate of change is not a constant rate of change, and cannot be used as an estimate of the rate of change except at or around the specific input value. This question is designed to get at students' across-time understanding of instantaneous rates of change.

Data Analysis

I examined the data from three angles, first comparing the two problems about the one-pound increase in weight (both linear and non-linear), then comparing the two problems about the twenty-pound increase in weight (both linear and non-linear), and finally comparing the two problems about the non-linear function (one-pound and twenty-pound).

For each comparison, I took an approach similar to Monk (1994) and created 2x2 contingency tables to display combinations of right or wrong answers. Students must have answered the question and provided some reasoning (not just "yes" or "no") in order to be included in the study. I performed McNemar's test (α= 0.05) to see if there were significant differences between the responses on the two questions. I also summarized the types of incorrect responses.

In order to classify the linear problem answer as correct (for both the one-pound and twenty-pound increases), students had to agree with the nurse's prediction and give reasoning that focused on the slope or constant rate of change being 2 mg per pound. For example, "Yes, I agree with the nurse because the slope is 2 mg per pound and it is a constant rate of change".

In order to classify the non-linear problem correct for the one-pound increase, students either had to agree with the nurse and give reasoning about using the derivative at 140 pounds is 2 mg per pound to estimate the increase, or to disagree with the nurse because the relationship is nonlinear and express uncertainty about using the derivative anywhere other than at 140 pounds. For example, "Yes, I agree with the nurse because the derivative at 140 pounds is 2 mg per pound, so she can use that to estimate the marginal change in dose from 140 pounds to 141 pounds."

In order to classify the non-linear problem as correct for the twenty-pound increase, students had to disagree with the nurse and give reasoning that focused on the non-linear nature of the model, and not being able to use the derivative to predict away from the input value of 140 pounds. For example, "No, I do not agree with the nurse because the derivative is the rate of change at 140 pounds only, and cannot be used to estimate the change in dose from 140 pounds to 160 pounds".

Findings

Students were much more successful answering the linear questions as compared to the non-linear questions (Figure 2), with 97% of students answering the linear one-pound question correctly and 95% answering the linear twenty-pound question correctly. This is compared to 63% of students answering the non-linear one-pound question correctly, and just 42% answering the non-linear twenty-pound question correctly.

	Linear		Non-Linear	
	1-pound	20-pound	1-pound	20-pound
% correct	97%	95%	63%	42%

Figure 2. Success rates of students in answering the linear and non-linear questions, N=74

Comparing Linear and Non-linear One-Pound Increases

To examine the results more closely, like Monk (1994), I coded students' answers so that combinations of rightness and wrongness could be examined, and presented the percentages in a 2x2 contingency table shown in Figure 3. The table shows the combinations of student responses for the two problems. For example, 62% of the students answered both the linear and non-linear questions correctly, and just 1% answered both of them incorrectly. And, 35% of the students answered the linear problem correctly but went on to answer the non-linear problem incorrectly. Similarly, 1% answered the linear problem incorrectly but went on to answer the non-linear problem correctly.

		Nonlinear		
Linear		Right	Wrong	Total
	Right	62%	35%	97%
	Wrong	1%	1%	2%
	Total	63%	36%	

Figure 3. Contingency table for correctness of answers to linear and non-linear change in dose for weight changes from 140 to 141 pounds questions, N=74

I performed McNemar's test to investigate the null hypothesis that the probability of getting the linear problem correct is the same as the probability of getting the nonlinear problem correct, and concluded that there was a significant difference in the results ($p < 0.0001$).

Focusing on the 26 students (35%) who answered the linear problem correctly but went on to answer the nonlinear problem incorrectly, 13 answered yes, that they agreed with the nurse and used some language about constant rates of change. For example, one answered, "I do agree. The derivative of a function shows the rate of change, which is constant in this case." And another answered, "Yes, she's simply looking at the rate of change if the weight is 140 (it's the same for all weights, a linear derivative". These students did not display an understanding of how the non-linear function comes into play, and instead relied on similar language as they had for the constant rate of change in the linear function.

Three students answered that they did not agree with the nurse, and gave the reason that they needed to know the slope at 141, or needed to know D'(140). For example, one student stated, "I don't agree. She needs to solve the derivative equation at D'(141) because D(w) is nonlinear". This is interesting because they recognized that the non-linearity of the function comes into play, but then they thought that they just needed the derivative at another point to answer the question.

In addition to these categories, one student answered "Not enough info to answer" and nine answered nonsensical answers, such as "No, the dosage would be 39,762 mg."

Comparing Linear and Non-linear Twenty-Pound Increases

I performed the McNemar's test on the 2x2 contingency table (Figure 4) to test the null hypothesis that the probability of getting the linear problem correct is the same as the probability of getting the nonlinear problem correct, and concluded that there was a significant difference in the results ($p < 0.0001$).

		Nonlinear		
Linear		Right	Wrong	Total
	Right	38%	57%	95%
	Wrong	4%	1%	5%
	Total	42%	58%	100%

Figure 4. Contingency table for correctness of answers to linear and non-linear change in dose for a 20-pound increase in weight, N=74

Focusing on the 42 students (57%) who answered the linear problem correctly but went on to answer the nonlinear problem incorrectly, 28 answered yes, that they agreed with the nurse because of a "constant slope" or "same slope." For example, one answered "Yes, because the rate of change dD/dw is 2 mg per pound," and another answered "Yes, because d-prime is the slope and it is constant."

Three answered that they did not agree with the nurse because they needed to know D'(160). For example, one student stated, "No, because the dosage increases by the derivative of 150 which will be different than 2 since it is nonlinear." These responses are similar to those who, in the one-pound problem, answered that they needed to know D'(141).

Two realized that it was not an appropriate reasoning, but concluded that because the relationship was nonlinear, the true dosage would be higher. They equated nonlinear with "exponential" or at least, increasing at an increasing rate.

For the remaining nine, one said "Not enough information to answer," one said it was "nonlinear but it was OK to make the nurse' prediction," one said "not entirely sure," and six gave nonsensical explanations or calculations.

Comparing Non-Linear One-Pound vs. Twenty-Pound Increases

Lastly, I performed the McNemar's test on the 2x2 contingency table (Figure 5) to test the null hypothesis that the probability of getting the linear problem correct is the same as the probability of getting the nonlinear problem correct, and concluded that there was a significant difference in the results ($p < 0.001$).

		20 pound increase in weight		
		Right	Wrong	Total
1 pound increase in weight	Right	41%	23%	64%
	Wrong	1%	35%	36%
	Total	42%	58%	100%

Figure 5. Contingency table for correctness of answers non-linear one-pound increase and non-linear twenty-pound increase, N=74

Focusing on the 17 students (23%) who answered the one-pound question correct but went on to answer the twenty-pound problem incorrectly, 13 answered that they agreed with the nurse because the slope stays the same. Even though they gave what looked like a reasonable answer for the one-pound increase, they went on to use the idea of constant rate of change for the twenty-pound increase. For example, for the one-pound correct answer, one student stated "Yes, I agree, because at point 140, D(w) is changing at 2 milligrams per pound," but then went on for the twenty-pound problem to say "Yes, I agree because the nurse calculated 340 mg using the slope of 2 mg per 1 pound."

Two students stated that they needed D'(160) to answer the problem, once again realizing that the non-linearity comes into play, but that what is needed is just the instantaneous rate of change at a different point (the "end" point).

One student equated non-linear with exponential, and said that the true dosage would be much higher than the nurse's estimate. One student said that even though it was non-linear, the nurse's prediction was valid. And lastly, one student answered that he agreed with the nurse but he was not sure why.

Summary of Findings

Three interesting themes emerged during the data analysis. First, students were successful at interpreting constant rates of change with linear relationships, with over 95% success on the two linear problems. Students had difficulties, however, interpreting the derivative in the context of the problem, what it represents, and how it can be used for approximations, with just 63% getting the one-pound increase correct and 42% getting the twenty-pound increase correct.

Secondly, a few students recognized the non-linear nature of the relationships, but went on to conclude that a derivative (just not the one given) was all that was needed to estimate the change in the dependent variable. This was especially noteworthy with the twenty-pound increase, where three students thought that the derivative at 160 pounds (instead of 140 pounds) was the missing piece of information.

Thirdly, a few students equated an increasing non-linear function with "exponential" or at least with "concave up." While not directly related to my research questions, it does raise questions as to how students view a "non-linear function."

Conclusions and Implications

General Conclusions

My research focused on student understanding of concepts that required an across-time understanding of functions (Monk, 1988). Specifically, students seem to have a solid understanding of the use of the slope of a linear equation to predict unit change. For example, 97% and 95% correctly agreed with the nurse's prediction of the increase in dosage using the linear model for both one-pound and twenty-pound increases, respectively. And, even though 64% also predicted the change in dosage for a one-pound weight increase correctly for the nonlinear function using the derivative at 140 pounds, more research in the form of student interviews should be performed to get at whether the correct answer stems from an understanding of the instantaneous rate of change as a measure of the marginal change, or because of a misconception about interpreting the derivative as a constant rate of change. Many of the students went on later to use the fact that $D'(140) = 2$ to agree with the nurse's prediction of the dosage for a 160-pound patient. This makes me question whether students were using this wrong interpretation to answer the question about the change in dosage 140 to 141 pounds. Interviews could uncover this more.

For students' understanding of the validity in making predictions from slopes and derivative, the results point to misunderstandings on the part of calculus students. For the twenty-pound increase problem, while 95% answered the linear problem correctly, only 42% reached the correct conclusion on the nonlinear problem. While more research in the form of interviews should be done, one can preliminarily conclude that students do not have a full understanding of the meaning of an instantaneous rate of change and an across-time understanding of the derivative as a function.

Depending on the crafting of the question, students sometimes give correct answers for wrong reasons, which makes it difficult to detect misconceptions (Bezuidenhout, 1998). This might be the case for the students whose answers I coded as "correct" for the non-linear one-pound increase. While they seemed to give correct reasoning, such a large number got the twenty-pound increase incorrect that I am unsure that they understood how the non-linearity came into play.

Bezuidenhout (1992) found that only 2% of participants were able to interpret the meaning of a derivative in the context of a problem. While my results are definitely more promising, more research is necessary to fully understand the actual student thinking surrounding the non-linear responses. Additional questions and interviews could get at student understanding of the instantaneous rate of change as an estimate of the marginal change, as well as when and how a derivative should be used to make predictions.

The struggles students had are similar to the struggles Monk found in his research surrounding students' across-time understanding of functions. Like Monk (1988), I suggest a separate study that would include an analysis through interviews of the mental processes of students who do not get the correct answers to these problems.

Revisiting the Research Questions

Thinking back on my two original research questions:

- Is there a relationship between calculus student understanding of slope and their understanding of derivative? Specifically, do students' abilities to interpret the slope as a constant rate of change make them more likely to be able to interpret the derivative as an instantaneous rate of change?
- Do students correctly use the slope and derivative to make valid predictions from models?

Students must have a clear understanding of a constant rate in order to understand instantaneous rate of change (Hackworth, 1994). This study supports this claim, as there was only one student who was able to successfully answer the non-linear derivative question after answering the constant rate of change question incorrectly. Rates of change in general must be understood by students to succeed in calculus (Hackworth, 1994), but this research adds to the set of findings that show that rates of change are not well-understood by first-year students, many of whom may have fundamental misconceptions (Bezuidenhout, 1998).

Without a solid of understanding of rates of change, students are not able to correctly use the slope and derivative to make valid predictions. In fact, almost two-thirds of the students incorrectly used the instantaneous rate of change as a constant rate of change, pointing out some of the misconceptions about the derivative's meaning as an instantaneous rate of change, and incomplete across-time understanding of the derivative as a function.

Future Research

I suggest more research be done in the form of interviews to understand student thinking about the instantaneous rate of change at a point, and why many used the instantaneous rate of change as a constant to make predictions. These interview questions would focus on student thinking about their answers to the questions about using the slope and derivative to make predictions.

I also suggest some changes to the survey if it were to be used again. One slight change would be instead of asking students whether they agree with the nurse's reasoning to ask instead how much confidence they have in the nurse's reasoning. I think this would uncover the students who I coded as correct for non-linear one-pound increase who do not understand the derivative as an estimate of the marginal change.

Teaching Implications

It is evident from previous research (Hackworth, 1994, Orton, 1984) that students lack a solid understanding of rates of change in general, something needed in order to understand slope as a constant rate of change and derivatives as instantaneous rates of change. This study expands on those previous findings by showing that students do not have a full understanding of the difference between a constant rate of change and an instantaneous rate of change.

How can we provide students with the general understanding of rates of change early in their mathematical career? At the early grades, more focus must be given on rates of change, even before the context of linear functions and slope. Students must learn to interpret rates of change in the context of questions, and understand how the rates of change affect the variables of interest.

What can we do as calculus college instructors? It is important to assess where our students are at coming into calculus, and provide the basic instruction of rates of change to fill in any gaps in knowledge. It is also important to make sure our students are able to answer rates of change questions that require across-time knowledge of functions, not the point-wise questions that are often in textbooks. If we want our students to understand the differences between instantaneous rates of change and constant rates of change, we have to ask them across-time questions that require them to reason and communicate about these difficult concepts.

References

Asiala, M., Cottrill, J., Dubinsky, E., & Schwingendorf, K. E. (1997). The development of students' graphical understanding of the derivative. *The Journal of Mathematical Behavior*, *16*(4), 399-431.

Barr, G. (1980). Graphs, Gradients and Intercepts. *Mathematics in School*, *9*(1), 5-6.

Barr, G., (1981). Some student ideas on the concept of gradient. *Mathematics in School,10(1),* 14-17.

Berry, J. S., & Nyman, M. A. (2003). Promoting students' graphical understanding of the calculus. *The Journal of Mathematical Behavior*, *22*(4), 479-495.

Bezuidenhout, J. (1998). First-year university students' understanding of rate of change. *International Journal of Mathematical Education in Science and Technology*, *29*(3), 389-399.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 352-378.

Ferrini-Mundy, J., & Graham, K. (1994). Research in Calculus Learning: Understanding of Limits, Derivatives, and Integrals. In J. J. Kaput & E. Dubinsky (Eds.), *Research Issues in Undergraduate Mathematics Learning: Preliminary Analysis and Results* (Vol. 33, pp. 29-45). Washington, DC: The Mathematical Association of America.

Franzosa, R. & Tyne, J. (2010). *Algebraic Models in Our World*. Dubuque, IA: Kendall Hunt

Hackworth, J. A. (1994). *Calculus Students' Understanding of Rate* (Doctoral dissertation, San Diego State University).

Holdren, J. P., & Lander, E. S. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, D.C.

Hughes-Hallett, D., McCallum, W., Gleason, A., Connally, E., Lomen, D., Quinney, D., …Tucker, T., (2013) *Calculus 6th Edition*. Hoboken, NJ: Wiley

Monk, G. S. (1994). Students' Understanding of Functions in Calculus Courses. *Humanistic Mathematics Network Journal*, *9*, 21-27.

Orton, A. (1983). Students' understanding of differentiation. *Educational Studies in Mathematics*, *14*(3), 235-250.

Orton, A. (1984). Understanding rate of change. *Mathematics in school*, *13*(5), 23-26.

Schoenfeld, A. H. (Ed.). (1987). *Cognitive science and mathematics education*. Lawrence Erlbaum.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences* (pp. 115-116). Boulder, CO: Westview Press.

Siegler, R. S. (2003). Implications of cognitive science research for mathematics education. *A research companion to principles and standards for school mathematics*, 219-233.

Stump, S. L. (2001). High school precalculus students' understanding of slope as measure. *School Science and Mathematics*, *101*(2), 81-89.

Thompson, P. W. (1994). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational studies in mathematics*, *26*(2), 229-274.

Ubuz, B. (2007). Interpreting a graph and constructing its derivative graph: Stability and change in students' conceptions. *International Journal of Mathematical Education in Science and Technology*, *38*(5), 609-637.

White, P., & Mitchelmore, M. (1996). Conceptual knowledge in introductory calculus. *Journal for Research in Mathematics Education*, 79-95.

Zandieh, M. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. *Research in Collegiate Mathematics Education. IV. CBMS Issues in Mathematics Education*, 103-127.

AN ORIGIN OF PRESCRIPTIONS FOR OUR MATHEMATICAL REASONING

Yusuke Uegatani
Research Fellow of the Japan Society for the Promotion of Science (Hiroshima University)

To build a supplementary theory from which we can derive a practical way of fostering inquiring minds in mathematics, this paper proposes a theoretical perspective that is compatible with existing ideas in mathematics education (radical constructivism, social constructivism, APOS theory, David Tall's framework, the framework of embodied cognition, new materialist ontologies). We focus on the fact that descriptive and prescriptive statements can be treated simultaneously, and consider both descriptive and instantiated models in our minds. This indicates that descriptive statements in mathematics come from our descriptions of models, and prescriptive statements come from the instantiatedness of the instantiated models and non-existence of counterexample. As a practical suggestion from the proposed perspective, we point out that careful communication is needed so that students do not recognize the refutation of their arguments as a denial of their way of mathematical thinking.

Key words: Inquiring minds, Prescriptive perspective, Mathematical reasoning

Introduction

Some undergraduate students seem to have only inadequately *inquiring minds* in mathematics, though inquiring minds are vital in continuing to study advanced mathematics. For example, a mathematician interviewed by Weber (2012) says, "I find students read a proof like they would read a newspaper and it's impossible to understand proofs that way" (p. 475). This comment implies that some students tend to accept proofs even before reading them, and as a result tend not to obtain the new insights that they would acquire through reading them. This is regarded as a "lack of inquiring minds" in this paper. In addition, these students also seem to uncritically accept most mathematical statements provided by their teachers in mathematics lectures. Following the distinction between a *mathematical attitude* and *an attitude toward mathematics* (Freudenthal, 1981, pp. 142–143), they lack mathematical attitudes, though they may have attitudes toward mathematics.

This lack of inquiring minds or mathematical attitudes may also be conceptualized as a lack of "mathematical integrity," a quality that involves commitment to mathematical truth (DeBellis & Goldin, 2006). To be specific, the phenomenon results from a lack of the unconscious belief that the discoverability of new mathematical results or the rediscoverability of already known mathematical results is open to everyone. For students without this *discoverability belief*, reading proofs or participating in mathematics lectures is not a process of (re)discovering mathematical results, but may instead be just a matter of encountering claims dependent on historical contingency, temporary human discourse, or authority. Some such students fail to make sense of mathematical statements, while others try to construct meanings such that the statements make sense to them. They do not check the validity of the statements, because they think that the statements are always correct if they only make sense. These students have difficulty in continuing to study mathematics. Therefore, in order to obtain practical implications from these cases to support successful mathematics learning, we need to identify the origins of discoverability beliefs and understand how they influence students.

For this purpose, it is not enough to explain the origin of one's discoverability belief as one's successful experience in discovering some mathematical results by oneself. As an example of such a successful experience, we may take a kind of sudden insight within problem solving, known as an *AHA! experience* (cf. Liljedahl, 2005). Discoverability beliefs also seem to depend on experiences of such subjective feelings. However, this explanation does not clarify the reasons why some students feel as if they have discovered something mathematical and others in the same lecture room do not. We may perhaps ascribe discoverability beliefs to uncontrollable subjective factors, but such a theoretical perspective is not useful for educational practice. For example, in terms of AHA! experiences, Liljedahl (2005) pointed out that "the environment for such an experience can be orchestrated, but the experience itself cannot" (p. 232). This implies that in order to get practical implications for the establishment of an adequate learning environment, we need to identify controllable objective factors that increase one's probability of a successful experience or decrease one's probability of an unsuccessful experience in (re)discovering some mathematical result.

One possible approach to this problem is the *epistemological* one. Although new mathematical findings may sometimes depend on empirical evidence, establishing the validity of mathematical statements does not need empirical support in many cases. What is needed to establish mathematical truth is usually just mathematical reasoning. Thus, one's feeling about discovery mainly depends on one's own process of establishing mathematical knowledge. An origin of discoverability beliefs can be supposed to consist in such an epistemological process of human mathematical reasoning if mathematical truth does not depend on arbitrary human judgments.

Several epistemological approaches to the process of establishing mathematical knowledge exist in mathematics education research, such as radical constructivism (Thompson, 2000; von Glasersfeld, 1995), social constructivism (Ernest, 1991, 1998), APOS theory (Dubinsky & McDonald, 2002), the three worlds of mathematics (Tall, 2004, 2008, 2011), embodied cognition (Lakoff & Núñez, 2000), and new materialist ontologies (de Freitas & Sinclair, 2013). However, none of these explain how the discoverability belief, or whatever its counterpart is in each theory, arises. (Of course, they do provide explanations for broader educational phenomena, and their scant attention to discoverability belief is thus forgivable, because each theoretical perspective has its own purpose.)

Thus, in order to obtain practical implications to support successful mathematics learning, we need a new supplementary theoretical perspective. As Cobb (2007) argued, "we should view the various co-existing perspectives as sources of ideas to be adapted to our purposes"; therefore, if the existing paradigms do not provide a direct solution, we must build a supplementary perspective integrating useful pieces of existing theoretical knowledge for a certain educational purpose. This paper attempts to build such a supplementary theory from which we can derive practical implications for the establishment of a learning environment where students can eventually acquire inquiring minds.

Sufficient Conditions of the Supplementary Theory

It is important to declare what conditions of the supplementary theory will be sufficient before trying to establish that supplementary theory. This will provide us with the needed constraints on the establishing process. In this regard, we make five assumptions in this paper.

The first assumption is that the objective factors determine uniquely the set of viable subjective knowledge. In this paper, we must identify controllable objective factors related to discoverability beliefs in the subjective processes of students' mathematical reasoning. Some readers may feel that this attempt is paradoxical because of the attempt to find objective factors in subjective processes. However, this paradox disappears if we specify that we are using the term "objective" to describe something from the observer's (e.g., the teacher's or the researcher's) perspective, and the term "subjective" to describe something from the learner's perspective. In radical constructivist theory, subjective knowledge is to an objective problem what a key is to a lock (von Glasersfeld & Cobb, 1984). Although no single particular key can be uniquely determined by the particular lock, the *set* of usable keys is physically uniquely determined by the lock. Similarly, although valid subjective knowledge appropriate to solving an objective problem cannot be uniquely determined, we can assume that some set of viable subjective knowledge is uniquely determined by the problem. We thus establish the possibility of identifying influential objective factors for the viability of subjective knowledge, and will try to build a theoretical framework to capture such factors.

The second assumption is that as a result of students' use of learning strategies, their cognitive development follows David Tall's theory (Tall, 2008) when they construct new mathematical concepts, even when reading proofs or participating in lectures. The theory partially incorporates APOS Theory (Dubinsky & McDonald, 2002) and conceptual metaphor theory (Lakoff & Núñez, 2000). It explains students' cognitive transition from the earliest pre-school mathematics to graduate mathematics. However, it mainly explains successful development (outcomes), and is not directly suggestive for affective aspects such as attitudes or beliefs. The framework necessary for our purpose will be one which explains how some students autonomously begin to use successful learning strategies, resulting in the kind of cognitive development described by Tall's theory, when reading proofs or participating in lectures. This explanation will elaborate an origin for discoverability beliefs.

The third assumption is that we can compare the degrees of freedom of the solutions to a certain objective problem with those of an equivalent problem. Following the first assumption, for any problem for any student, the set of viable subjective knowledge for solving the problem is unique, but we cannot predict which knowledge in the set will actually survive or vie for survival, because many accidental factors influence the student. On the other hand, even if two problems are objectively equivalent, it is not necessarily warranted that two problems are subjectively similar to each other. If this third assumption holds, we can choose one among the equivalent problems, which will increases the probability that the intended knowledge actually vies. This paper will build a theory satisfying the third assumption.

The fourth assumption is that the patterns of mathematical reasoning are common among students but that their consequences can differ because accidental factors cause students to arbitrarily arrange the patterns in their reasoning processes. This assumption is, albeit indirectly, supported by the existing research. For example, Nesher (1987) indicates that "most [misconceptions] are overgeneralizations of previously learned, limited knowledge which is now wrongly applied" (p. 37). Even unsuccessful students with misconceptions, as well as successful students, have some mathematical attitude toward generalization of their subjective knowledge. Another example is from the research on concept images. According to Tall and Vinner (1981), students have their own subjective images of each concept; some students successfully use concept images (Pinto & Tall, 2002) and others use them unsuccessfully (Tall & Vinner, 1981; Vinner & Dreyfus, 1989). However, both successful and unsuccessful mathematical thinking have some aspects in common. As a radical extrapolation from this fact, we make the forth assumption: if some pattern of mathematical reasoning is appropriately modeled, the process of loss of discoverability beliefs related to it can be explained in terms of the following four steps. First, all students use common patterns of mathematical reasoning in early learning. Second, however, due to accidental factors, some students fail to learn mathematics, in spite of the fact that their learning strategies are the same as those of successful students. Third, unsuccessful students mistakenly perceive that the reason why they failed is because they are using inadequate learning strategies. Finally, as a result, they eventually lose their discoverability beliefs and do not come to use successful learning strategies. Therefore, our theoretical framework must provide an appropriate model of human mathematical reasoning. One of the practical goals of the framework will be to help students correctly recognize the validity of their initial learning strategies, because it is difficult to completely remove accidental factors.

The final assumption is that successful experience of mathematical discovery depends mainly on mathematical reasoning, though some types of mathematical discovery may depend on physical evidence. If this assumption does not hold, we will not be able to understand why the validity of mathematical knowledge depends mainly on reasoning. On the other hand, this assumption implies that mathematical reasoning must have some prescriptive aspects. One example of such is that if the propositions $P \rightarrow Q$ and P are true, then the proposition Q should be true. If students do not perceive this prescriptive proposition from their mathematical reasoning, then they will not have experienced mathematical discovery.

On the basis of these five assumptions, the main research task of this paper is to model the common patterns of human mathematical reasoning. The model must satisfy the following four sufficient conditions. First, it must identify the factors influencing the degrees of freedom of the solutions to a problem. Second, the model must explain how higher degrees of freedom tend to produce more accidental factors. Third, the model shows that the mechanisms of both successful and unsuccessful reasoning are the same except for the tendency to accept the influence of accidental factors. Finally, the model explains that a certain type of arrangement of reasoning patterns causes students to feel the presence of prescriptiveness in the knowledge at stake.

In the following section, we will discuss the dual aspects of mathematical reasoning: *prescription* and *description*. Through the elaboration of both aspects, we will eventually succeed in modeling a mathematical reasoning that can satisfy the above conditions.

Duality of Prescription and Description

Ernest (1998) pointed out the limitations of *prescriptive* accounts of mathematics:

> Absolutist philosophies of mathematics such as logicism, formalism, and intuitionism attempt to provide *prescriptive* accounts of the nature of mathematics. Such accounts are programmatic, legislating how mathematics should be understood, rather than providing accurately *descriptive* accounts of the nature of mathematics. Thus they are failing to account for mathematics as it is, in the hope of fulfilling their vision of how it should be. (pp. 50-51, italics in the original)

Thus, Ernest's (1998) social constructivism takes a descriptive stance. It provides no account of which way of doing mathematics is correct, but rather describes how people do mathematics. Other existing research perspectives for mathematics education also take descriptive stances. They provide no account of which method of understanding mathematics is correct, but merely explain how students do mathematics. However, the preceding discussion is based on the following implicit assumption: we must exclusively choose prescriptive or descriptive philosophies. Both the prescriptive statement "X should be Y" and the descriptive statement "X is Y" can be simultaneously correct.

For example, consider a group $(G,*)$. Suppose that G is a set, and that $*$ is a binary operation on G. The group axioms are as follows: (i) For all a, b in G, $a * b$ is also in G. (ii) For all a, b and c in G, $(a * b) * c = a * (b * c)$. (iii) There exists an element e in G such that, for every element a in G, the equation $a * e = e * a = a$ holds. (iv) For each a in G, there exists an element b in G such that $a * b = b * a = e$, where e is the element defined in axiom (iii). From these axioms, we can derive the statement that the element e postulated in (iii) is unique, and we will say that e postulated in (iii) should be unique if someone argues that there are many elements postulated in (iii). In this case, both statements (involving "is" and "should be") appear correct. This is explained by distinguishing between *in* and *out of* the axiomatic system. The statement that the element e postulated in (iii) is unique is a description of components *in* the system. The statement that the element e postulated in (iii) should be unique (or, more strictly, the statement that we should argue that e postulated in (iii) is unique) is a prescription for us who are *out of* the system. It is important that the element e (or the entity in the system) is not itself bound by the rules of logic, but that all thinking subjects who are out of the system and agree on the group axioms have an obligation to obey some logical inference rules.

In general, a descriptive statement in an axiomatic system and the corresponding prescriptive statement out of the system can be simultaneously correct, because we can always distinguish between in and out of the given system. It is, therefore, an unjustifiable assumption that we cannot simultaneously consider both prescription and description. If we have the ability to self-reflect, and to distinguish between the outside of an axiomatic system and the overall framework that contains the inside and the outside of the system, then prescriptive statements and descriptive statements are dual properties of the overall framework. In addition, it is also important that humans out of the system are prescribed, and the entities in the system are described at the same time.

Origin of Prescription

If our reasoning always followed the rules of formal logic, the discoverability belief would be justified by the independence between these rules and human minds. In general, it is difficult to describe the actual practices of mathematics only by formal logic (e.g., Fallis, 2003). Thus, we argue that the schemata of descriptions actually prescribe human reasoning.

The schema of descriptions is, for example, the format of implication statements "$P \to Q$." We do not assume that it pre-existed the modus ponens. Rather, we argue that modus ponens pre-existed the schema "$P \to Q$," and that the schema was invented to describe a situation where one may infer Q after knowing that P is true. Given the propositions P and $P \to Q$, we usually deduce proposition Q for any propositions P and Q. This does not imply the validity of modus ponens, but implies that there can be a situation where one may infer Q after knowing that P is true. Similarly, the rule of conjecture elimination (inferring P from $P \land Q$) pre-existed the schema "$P \land Q$," and the rule of universal instantiation (inferring $A(a)$ for any element a from $\forall x\, A(x)$) pre-existed the schema "$\forall x\, A(x)$." In general, an inference rule pre-existed its related schema. Thus, what one should infer depends on how one describes a given situation, and not on formal logic.

From this perspective, it is necessary to identify what determines a valid description of the situation. Next, we shift to the question of how descriptive statements arise.

Origin of Description

In mathematics, some descriptive statements are contained within the axioms of the system under consideration, but even in advanced mathematics, we do not always think in completely formalized systems. We propose that, instead, descriptive statements originate from *models* in our minds. In the present paper, the term *model* has a dual meaning. In this regard, Mason's (1989) idea is highly suggestive. According to Mason (1989), mathematical abstraction is described as "a delicate shift of attention from seeing an expression *as* an expression of generality, to seeing the expression *as* an object or property" (p. 2, italics in the original). Using the idea of "a shift of attention," we will show the dual meaning of "model."

One meaning is "something that a copy can be based on because it is an … example of its type" ("Model," n.d.-a). We call this an *instantiated model*. For example, the set of all integers, together with the operation $+$, is an instantiated model of a group in our minds, because it is a typical example of a group. With this in our minds, we can easily understand any example of a group by analogy. We can also show that the set of all integers with the operation $+$ is an instantiated model satisfying the group axioms. Similarly, because the experience of typicality can depend on subjective experiences, any example of a group can be an instantiated model. As it has not only the essential features of a group, but also non-essential ones, it has more information than a group as an abstract object without any non-essential features of a group. In general, an instantiated model satisfies a certain set of axioms, and carries more information than an abstract object without any properties which the axioms do not imply. A set of axioms do not have to be commonly accepted. Arbitrary logical expressions may be axioms. If a set of axioms is consistent, there exists at least one instantiated model for them.

Another meaning of the term "model" is "something that represents another thing … as a simple description that can be used in calculations" ("Model," n.d.-b). We call this a *descriptive model*. For example, a line in mathematics may be regarded as a descriptive model of a physical line, such as that made by a pencil, in our minds. A line in mathematics is defined by focusing attention on only some of the features of a physical line. It is a result of neglecting uninteresting features that. While a physical line does have width, we usually require in mathematics that a line have no width. In general, a descriptive model is created by focusing attention on only some of the features of other descriptive models or physical objects. Such a temporal creation is then refined with certain provisos (e.g., "it has no width"). The provisos prevent us from focusing attention on uninteresting features of the source descriptive models or objects.

Most relevant here is the relativity between instantiatedness and descriptiveness. That is, when we focus attention on some essential features of an instantiated model, the abstract object constrained by the logical expressions of those features is a descriptive model of the instantiated model. When we create a new object by adding some extra features to an abstract object that is a descriptive model, the new object is an instantiated model of the descriptive model. In other words, any model in our minds can always be both instantiated and descriptive. Any model other than a physical object is an instantiated model of more abstract models or objects, and it is simultaneously a descriptive model of more concrete models or objects. The relativity between instantiatedness and descriptiveness allows us to dispense with the distinction between the terms "model" and "object." In this sense, both terms may be used interchangeably, because every model can become an object of thought, and vice versa.

By using the term "model," one of the predominant origins of descriptive statements in mathematics can be explained as descriptions of models in our minds. We will provide two examples: the fundamental theorem of cyclic groups, and the construction of an equilateral triangle. Let us explain their possible models, for example, in the author's mind.

The fundamental theorem of cyclic groups: The theorem states that every subgroup of a cyclic group is cyclic. Let $\langle g \rangle$ be a cyclic group generated by g. Following the definition of a cyclic group, $\langle g \rangle$ simply consists of $\cdots, g^{-2}, g^{-1}, e, g, g^2, \cdots$; there is no other element in $\langle g \rangle$. If a subgroup of $\langle g \rangle$ has n different elements, they can be represented by $g^{k_1}, g^{k_2}, \cdots, g^{k_n}$. From the group axioms, the subgroup contains $g^{\mathrm{GCD}(k_1,k_2,\cdots,k_n)}$, and $g^{\mathrm{GCD}(k_1,k_2,\cdots,k_n)}$ generates all elements in the subgroup. Thus, the theorem seems to be true.

This way of creating descriptions of models in our minds implies various prescriptions. For example, when someone says that $\langle g \rangle$ might not contain e, the author should argue that $\langle g \rangle$ always contains e because $\langle g \rangle$ is an example of a group. As another example, when someone points out that the order of a subgroup of $\langle g \rangle$ is not always finite, the author should recognize that an example of a subgroup of $\langle g \rangle$ in his mind is too specific.

The construction of an equilateral triangle on a given line segment: Let AB be the given line segment. Draw a semicircle with center A and radius AB. Again, draw a semicircle with center B and radius BA on the same side as the first semicircle. Let C be the point of intersection of the semicircles. Then, the triangle ABC is equilateral. This is because the semicircles centered at A and B have radii of equal length, and all three segments AB, BC, and CA are the length of their radii. Thus, the construction seems to be valid.

There are also various prescriptions in this case. For example, when someone says that the three edges AB, BC, and CA are not always equal, the author should argue that they are always equal, for the following reason. The point C is regarded as our instantiated model of the points on the semicircles A and B; the pairs CA, AB and AB, BC are regarded as our instantiated models of equivalent radii, and the lengths of AB, BC, and CA are regarded as our instantiated models of the transitivity rule. As another example, if someone points out that the author's consideration depends on the belief that the two semicircles always intersect with each other, he should recognize that his consideration depends on a visual representation.

Generally speaking, descriptive statements of some mathematical objects are created by accessing their models in human minds, and then describing these models. Given an axiomatic system (that is, a descriptive model), one creates an instantiated model of the given descriptive model in mind. Creating a descriptive statement in the system is creating a descriptive model of the current model in mind. There are two types of creation. One creates a description of a common property among all the instantiated models of the given descriptive model. The other creates a description of a property satisfied by only a particular instantiated model of the given descriptive model. If one mistakenly argues something based on the latter type, and someone points this out, then one should recognize the mistake (for example, that an example of a subgroup of $\langle g \rangle$ is too specific, or the consideration of an equilateral triangle depends on a visual representation). Descriptive statements in mathematics, therefore, can come from descriptions of models in our minds, and prescriptive statements can come from the instantiatedness of the instantiated models and non-existence of counterexamples. From this perspective, the reason why proofs and refutations (Lakatos, 1976) occur in the history of mathematics might be because humans (including mathematicians) sometimes create a description of a property satisfied by only a particular instantiated model of the given descriptive model.

Conclusion

In order to obtain the practical implications to support successful mathematics learning, especially with regard to discoverability, the author attempts to build a model of mathematical reasoning from a new theoretical perspective, presupposing the presence of mental models in the human mind. This paper asserts that strictly two types of mathematical reasoning exist, involving either the creation of instantiated or descriptive models from the mind's present model.

This proposed model of mathematical reasoning satisfies the four conditions presented in the second section. First, a factor influencing the degree of freedom in solving a problem lacks sufficient constraints to ignore non-essential features. For example, student overgeneralization of certain mathematical topics can be attributed to the creation of descriptive models focusing on their non-essential features. In other cases, student misjudgment might be attributable to an overly specific mathematical concept image caused by the creation of instantiated models with additional non-essential features.

Second, if a learning environment permits students to focus on non-essential features, the probability of invalid mathematical reasoning will increase. The third vital property for successful reasoning entails focusing solely on the essential features that educators wish to teach; in contrast, unsuccessful reasoning is typified by a focus on non-essential features.

Finally, students who focus entirely on essential features will feel a sense of prescriptiveness. If a descriptive model of a common property among all instantiated models of a current mental model is created, prescriptiveness arises from subjective non-falsifiability. Discoverability beliefs originate from the repeated exposure of non-existent counterexamples.

As a practical suggestion from the proposed perspective, we point out that students might lose the discoverability belief if they recognize the refutation of their argument as a denial of their way of mathematical thinking. What the refutation actually denies might not be their attitude toward creating an instantiated model of the given descriptive model, but only the particular instantiated model contingently created at that time. If creating an instantiated model and describing it is an essential process of mathematics, a chain of reasoning means a chain of creating instantiated models or descriptive models of the already-created models. Then, many chains of reasoning are not deductive. If a student seems to mistakenly make a non-deductive chain of reasoning, the teacher should carefully communicate with the student, and try to recognize which chain would make such a conclusion. Otherwise, proofs and refutations do not work well as a social construction of mathematical knowledge in classrooms, and intersubjectivity cannot be established. In particular, it seems to be important for the teacher to pay attention not only to the student's conclusion but also to their attitude toward developing new findings in order to foster inquiring minds in mathematics. This teacher's attention can be one of controllable objective factors that increase one's probability of a successful experience or decrease one's probability of an unsuccessful experience in (re)discovering some mathematical result.

There are at least two limitations of the proposed perspective. First, it is still not clear whether it is completely compatible with each existing research perspective. Second, the above practical suggestion is still based on assumptions whose validity is not always warranted (for example, whether reasoning always means creating models). The suggestion describes only a possible situation in classrooms. Further development of our theoretical framework in this regard provides an avenue for future research.

Acknowledgment

This work was supported by Grant-in-Aid for JSPS Fellows No. 252024. Any opinion stated here is that of the author and does not necessarily reflect the views of JSPS.

References

Cobb, P. (2007). Putting philosophy to work: Coping with multiple theoretical perspective. *Second Handbook of Research on Mathematics Teaching and Learning: a Project of the National Council of Teachers of Mathematics*, *1*, 3–38.

De Freitas, E., & Sinclair, N. (2013). New materialist ontologies in mathematics education: the body in/of mathematics. *Educational Studies in Mathematics*, *83*(3), 453–470.

DeBellis, V. A., & Goldin, G. A. (2006). Affect and Meta-Affect in Mathematical Problem Solving: a Representational Perspective. *Educational Studies in Mathematics*, *63*(2), 131–147.

Dubinsky, E., & McDonald, M. A. (2002). APOS: A Constructivist Theory of Learning in Undergraduate Mathematics Education Research. In D. Holton, M. Artigue, U. Kirchgräber, J. Hillel, M. Niss, & A. Schoenfeld (Eds.), *The Teaching and Learning of Mathematics at University Level* (Vol. 7, pp. 275–282). Dordrecht: Kluwer Academic Publishers.

Ernest, P. (1991). *The philosophy of mathematics education*. Falmer Press.

Ernest, P. (1998). *Social Constructivism as a Philosophy of Mathematics*. SUNY Press.

Fallis, D. (2003). Intentional gaps in mathematical proofs. *Synthese*, *134*(1), 45–69.

Freudenthal, H. (1981). Major problems of mathematics education. *Educational Studies in Mathematics*, *12*(2), 133–150.

Lakatos, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press.

Lakoff, G., & Núñez, R. E. (2000). *Where Mathematics Comes from: How the Embodied Mind Brings Mathematics Into Being*. Basic Books.

Liljedahl, P. G. (2005). Mathematical discovery and affect: the effect of AHA! experiences on undergraduate mathematics students. *International Journal of Mathematical Education in Science and Technology*, *36*(2-3), 219–234.

Mason, J. (1989). Mathematical abstraction as the result of a delicate shift of attention. *For the Learning of Mathematics*, *9*(2), 2–8.

Model. (n.d.-a). *Cambridge Dictionaries Online*. Cambridge University Press. Retrieved from http://dictionary.cambridge.org/dictionary/british/model_1

Model. (n.d.-b). *Cambridge Dictionaries Online*. Cambridge University Press. Retrieved from http://dictionary.cambridge.org/dictionary/british/model_4

Nesher, P. (1987). Towards an instructional theory: the role of student's misconceptions. *For the Learning of Mathematics*, *7*(3), 33–40.

Pinto, M., & Tall, D. (2002). Building formal mathematics on visual imagery: A case study and a theory. *For the Learning of Mathematics*, *22*(1), 2–10.

Tall, D. (2004). Building theories: The three worlds of mathematics. *For the Learning of Mathematics*, *24*(1), 29–32.

Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal*, *20*(2), 5–24.

Tall, D. (2011). Crystalline concepts in long-term mathematical invention and discovery. *For the Learning of Mathematics*, *31*(1), 3–8.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151-169.

Thompson, P. W. (2000). Radical Constructivism: Reflections and Directions. In L. P. Steffe & P. W. Thompson (Eds.), *Radical Constructivism in action: Building on the Pioneering Work of Ernst von Glasersfeld* (pp. 291–315). Routledge.

Vinner, S., & Dreyfus, T. (1989). Images and Definitions for the Concept of Function. *Journal for Research in Mathematics Education*, *20*(4), 356.

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. The Flamer Press.

Von Glasersfeld, E., & Cobb, P. (1984). Knowledge as environmental fit. *Man-Environment Systems*, *13*, 216–224.

Weber, K. (2012). Mathematicians' perspectives on their pedagogical practice with respect to proof. *International Journal of Mathematical Education in Science and Technology*, *43*(4), 463–482.

25

EXPLORING DIFFERENCES IN TEACHING PRACTICE WHEN TWO MATHEMATICS INSTRUCTORS ENACT THE SAME LESSON

Joseph F. Wagner
Xavier University

Karen Allen Keene
North Carolina State University

Investigating teacher practice at all educational levels has become an important research arena. We consider the teaching of inquiry-oriented differential equations in undergraduate classrooms by comparing two enactments of the same fragment of a student-centered curriculum by two mathematics professors. We highlight differences in the professors' practices and the consequent classroom results by analyzing the professors' participation in whole-class discussions and the decisions they made during class. By considering how the same written curriculum can be enacted in very different ways in undergraduate level mathematics classrooms, we call for greater attention to research on the relationships between a written curriculum, an enacted curriculum, and student outcomes.

Key words: Teaching Practice, Enacted Curriculum, Student-Centered Instruction

Science, technology, engineering, and mathematics (STEM) disciplines have been increasingly identified as a priority for educational improvement and innovation in the United States. The Department of Commerce (2012) listed mathematics and science education as one of six "alarms" that require our utmost attention in the 21st century. One way to improve STEM education is to improve student learning of mathematics at the university level. The mathematics education research community has begun to respond by creating new student-centered curricular materials for undergraduate mathematics courses that emphasize student inquiry, discovery, and problem solving. The effect of these efforts, however, depends not simply on the materials themselves, but on the ways that undergraduate instructors make use of them. To highlight this distinction, we examine the mathematics teaching of undergraduate instructors by interpreting their teaching practice as their participation in the *enactment* of a curriculum.

There is considerable research on teaching practice at the elementary and secondary level. Ball and Forzani (2009) indicated that study of "the work of teaching" is a particularly important area. However, the recent work in curriculum research at the K-12 level (i.e. Ridgway, Zawojewski, Hoover, & Lambdin, 2003; O'Donnell, 2008; Tarr, Reys, Reys, Chávez, Shih, & Osterlind, 2008) has not been extended to the university level. We believe that it is essential to study how a curriculum is enacted, especially as more innovative curricula are introduced at the college level. In this paper, we report on the results of our investigation of two university mathematicians' enactment of an innovative, student-centered curriculum. Because of the magnitude of the available data, we have chosen to focus only on the instructors' first day using the curriculum. Although much could be learned from the study of curriculum use over time, we believe that an analysis even of this single day supports the arguments we make here. In particular, we consider the following research questions:

- How might university professors differ in their enactments of (the first day of) an innovative, inquiry-oriented differential equations curriculum?
- How do such differences in curriculum enactment affect students' opportunities to learn?

Using classroom and interview data, we demonstrate how very different outcomes resulted from two instructors' use of the same curricular materials on their first day of class, particularly in terms of opportunities for student learning. We argue the need for RUME researchers to further research that does not isolate teachers' practice and curriculum enactment from student learning.

Literature Review and Theoretical Framework

Research on mathematics professors' teaching practice

After an extensive review of the research literature, Speer, Smith & Horvath (2010) concluded, "[R]esearch on collegiate teachers' actual classroom teaching practice is virtually nonexistent" (p. 99). Since that time, some RUME researchers have begun to respond. Attention has been given, for example, to teaching practice in lecture-based classrooms (Patterson, Thompson, & Taylor, 2011; Fukawa-Connelly, 2012; Trenholm, Alcock, & Robinson, 2012). A variety of studies investigating mathematics professors as they enact more student-centered curricula have also been published. (See, for example, *Journal of Mathematical Behavior*, 32, 2013). Even before Speer et al.'s (2010) call for research, Wagner, Speer, and Rossa (2007) reported on one instructor's knowledge as he implemented an inquiry oriented DE course. They identified forms of knowledge apart from mathematical content knowledge that are essential to reform-oriented teaching, and highlighted how knowledge acquired through more traditional instructional practices may fail to support research-based forms of student-centered teaching. Speer and Wagner (2009) considered the role of pedagogical content knowledge and mathematical knowledge for teaching through the construct of analytic scaffolding. Early analyses that led to this paper were first presented by Wagner (2007) in a preliminary report on the differences in instructors' use of inquiry-oriented curricular materials. Lee, Keene, Lee, Holstein, Ryals, and Eley (2008) suggested the construct of *mathematical content move* to discuss one mathematician's practice while first implementing an inquiry-oriented differential equations curriculum.

A few researchers have also responded to Speer et al.'s call for research *comparing* the teaching practice of instructors teaching the same course. Johnson, Caughman, Fredericks, and Gibson (2013) discussed their case studies of three mathematicians, finding three themes that emerged from interviews and reflections with these mathematicians: curriculum coverage; goals for student learning; and the role of the teacher. Pinto (2013) compared the lessons of two teaching assistants who individually interpreted and implemented the same lesson plan for a calculus class very differently. Our current work distinguishes itself, however, by contrasting the practices of two experienced mathematics instructors implementing identical curriculum materials. Additionally, we use the construct of *curriculum enactment* as a theoretical framework to help us better make sense of what happens in the classroom and how it affects student outcomes.

Research on K-12 curricula

In the 1990s, the National Science Foundation funded several new curricula to be used in mathematics teaching at the elementary, middle, and high school levels. These curricula came to be known as "reform curricula," and by 2000 there was significant interest in the success of these new curricula. This launched a myriad of studies and caused K-12 mathematics education researchers to think hard about the connections between a curriculum and student learning, as well as what it means for a curriculum to be "successful." Remillard and Bryans (2004) called attention to the significant yet largely unstudied role of the teacher in determining the way any curriculum is used to influence subsequent student learning:

> Focusing on the objectively given structures, we see that the resources provided in any curriculum represent a complex set of plans, activities, scripts, suggestions, information, explanation, and messages that have both textual and visual entailments and are likely to speak to different readers in different ways. We know little about how teachers engage these varied offerings. (p. 234)

Stein, Remillard and Smith (2007) developed a curriculum phase diagram to provide researchers with a framework for considering the transition of a curriculum from its written to its enacted form, and to subsequent student outcomes. (See Figure 1.) The *written curriculum* refers to all of the student materials, teacher materials, and supplemental materials that the curriculum authors originally designed. The *intended curriculum* includes the teacher's plans for the class and for how the materials will be implemented, and expectations of how students will engage with them. (In this paper, we will not address the intended curriculum directly, but it will be discussed peripherally in the results.) The *enacted curriculum* is what actually happens in class. We interpret the teacher's role in enacting the curriculum as what other mathematics educators often call teachers' *practice*. In other words, the enacted curriculum is determined largely (though not entirely) by the teacher's practice, including, but not limited to, the mathematical activities he actually uses, the decisions he makes during class, the questions he asks, and the time he allows for different tasks. Finally, the triangle at the right indicates the student learning that occurs as a result of the enacted curriculum.

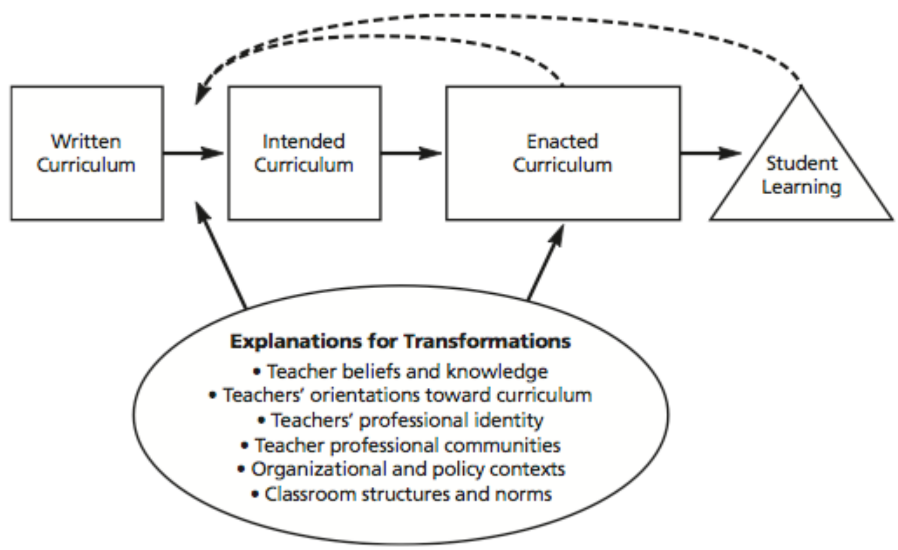


Figure 1: Temporal phases of curriculum use (Stein et al., 2007, p. 322)

The cloud below the curriculum enactment flow diagram indicates areas of investigation of what influences the transformation from the written to the enacted curriculum. Many of these issues are currently being studied in undergraduate mathematics education.

The term *curriculum* has not traditionally been used to describe the content and activities of undergraduate courses in mathematics. Typically, professors have chosen a textbook whose content and topics often do not vary much from one book to another, and traditional lecture approaches to teaching have not lent themselves to curricular analysis as the more extensive sets of books, activities, instructor notes, and support materials do in K-12 mathematics teaching have. The introduction of more student-centered curricular materials has brought the notion of implementing curriculum to the notice of university-level mathematics education researchers.

Because of this, RUME researchers may benefit from attending to what has been learned by research in K-12 education.

In her review of curriculum research, Remillard (2005) called attention to the variety of ways that educational researchers have either interpreted or simply assumed that curriculum implementers (in our case, mathematics professors) interact with a curriculum. Remillard identified four such interpretations or assumptions. First, and particularly in older studies, researchers have treated a curriculum as fixed and unchanging, as it is in its written form, and so the role of the teacher is merely to deliver its content. Within this perspective, it is assumed that it is possible for a teacher to implement a curriculum exactly as intended by its authors. Failure to remain "faithful" to the curriculum is essentially to subvert it. This approach leads to a positivist expectation that direct correlations can be made between the quality of a written curriculum and subsequent student progress. A second approach is to treat curriculum materials as but one of many sources available to a teacher in planning for instruction. This assumes that individual teachers have more agency in affecting the enacted curriculum, and that written curriculum materials provide a supportive role. A third approach perceives the teacher as an interpreter of the written curriculum, suggesting that fidelity between the written materials and the enacted curriculum is really not possible, as each individual instructor brings his or her own knowledge, experience, and values to bear in making use of the materials. Finally, a fourth approach treats instructors and curriculum materials as being in a dynamic and interactive relationship. In this understanding, the written materials exert an influence on the teacher, just as the teachers' own background influences his or her interpretation of how the materials may be used.

We are inclined to adopt this fourth perspective, avoiding a more positivist approach that might treat either or both of the mathematicians in this study as "failing" to implement the curriculum materials they were using. It seems to us that the materials given to them exerted a significant impact on their own understanding of how the curriculum might be enacted, but each instructor brought very different backgrounds that led to notably different interpretations of how the materials might be used. It is not evident to us that either "success" or "failure" is an appropriate interpretation of what occurred in their classrooms, nor do we believe that the standard of "faithfully implementing the curriculum" is a meaningful or measurable expectation to place on their shoulders.

The rise of inquiry-oriented curricula

In the past ten to fifteen years, a number of research-based sets of curricular materials have been written to support student-centered, and inquiry-oriented learning of undergraduate mathematics, including the NSF-supported Teaching Abstract Algebra for Understanding (http://www.web.pdx.edu/~slarsen/TAAFU/home.php), Inquiry-Oriented Linear Algebra (http://www.math.vt.edu/iola/index.php) and Inquiry-Oriented Differential Equations (Rasmussen, 2006). These materials have inspired a wide variety of research on both the teachers and students who have used them. Much of this research, we contend, may rightfully be placed within the cloud of "Explanations for Transformations" in Figure 1. That is, this research involves the study of phenomena that Stein et al. (2007) identified as directly influencing the *transformation* of a written curriculum into its eventual enactment. We note, however, that the role of these phenomena in effecting such transformations is rarely addressed by their researchers or the larger RUME research community. We believe that this risks fostering research on teaching practice that divorces it from its subsequent effect on student learning. The purpose of this paper is to call attention to this problem and encourage deeper reflection on some of our research priorities.

Methods

Data for the current study are taken from a much larger collection gathered as Prof. Gage and Prof. Paxton each taught a semester course in Differential Equations, two years apart, at a private, liberal arts university in the Midwest. The students in the class were primarily majors in mathematics or one of the physical sciences. Both instructors had doctorates in mathematics and each had been teaching for more than fifteen years at the university level. Both used the same set of curricular materials for an Inquiry-Oriented Differential Equations (IO-DE) course developed by Rasmussen (2006). Their previous experience involved more traditional lecture style courses, and neither had any particular exposure to, or training in, other methods of instruction. Each had, however, expressed some dissatisfaction with his prior student learning outcomes.

The IO-DE curriculum is a research-based set of curriculum materials including student tasks, instructor notes, and other instructional materials intended to support students as they "reinvent" graphical, symbolic, and numerical techniques to solve ordinary differential equations and linear systems of differential equations under the guidance of an instructor. The curriculum is designed to support student-centered learning, with activities cycling between small-group and whole-class investigations and discussions. The first activity presented to students on the first day of class is shown in Figure 1.

In this problem we study systems of rate of change equations designed to inform us about the future populations for two species that are either competitive or cooperative. Which system of rate of change equations below describes a pair of competitive species, and which system describes a pair of cooperative species? Explain your reasoning.

$$\frac{dx}{dt} = -5x + 2xy \qquad \frac{dx}{dt} = 4x - 2xy$$
$$\frac{dy}{dt} = -4y + 3xy \qquad \frac{dy}{dt} = 2y - xy$$

Figure 2: First activity of IO-DE curriculum

Almost all of each instructor's classes were videotaped with two cameras, one following the instructor and another focused on a selected small group of students. Audio-taped interviews were conducted with each instructor by the first author several times prior to the semester and after almost every class, with interviews varying between 15 and 75 minutes in length. For the present study, complete transcripts were made of the whole-class discussions for each instructor's first day of class, and significant portions of the interviews carried out near the first day of class were also transcribed.

The instructors' contributions to the whole-class conversations were coded using a coding scheme inspired by Wells and Arauz (2006) to determine the role that each turn of talk played in the conversation. The codes (some described in more detail below) were designed to capture the nature of each comment and each question, as a means highlighting the distinct conversational differences between the two classes. The two authors coded the transcripts independently using 18 possible codes, with 72% and 73% agreement for Prof. Gage's class and Prof. Paxton's class respectively. Disagreements were resolved by mutual discussion. Coding counts may be easily compared because both classes had segments of whole-class discussions that lasted nearly identical amounts of time (35:35 and 35:41 for Gage and Paxton, respectively). Because the focus of this analysis is on the instructors' practices, we do not code or otherwise attend directly to the students' contributions to the class conversations. In addition to making the analysis more tractable, we believe that a great deal is revealed even under this limitation.

Findings and Analysis

Analysis of professors' practice as seen in whole class discussion

Prof. Gage and Prof. Paxton each used the same written curriculum materials and student activities on their respective first days of class. Based on the interviews with the instructors before and after class, each instructor's intended curriculum gave every appearance of being closely aligned to the written curriculum. Even more, at first glance, the enactments of the two classes looked a great deal alike. In particular, we observed the following:

- Each instructor attempted to promote group learning and foster small-group and whole-class discussions.
- Each instructor expressed desire for students to develop their own knowledge, without dependence on the instructor's authority.
- Each instructor used open-ended questioning, and encouraged students to present their work and ideas to the class and to each other for critique.
- Each instructor made efforts to address both mathematical and social norms.

At closer consideration, however, significant differences could be observed between the conversation and character of the two classes. Prof. Gage and Prof. Paxton appeared to exercise different ways of directing and participating in the whole discussions. Table 1 provides excerpts of their contributions to a segment of the whole class discussion (with student contributions omitted) that highlight the differences that we found typical of the two classes. In particular, Prof. Gage's questioning appeared to be very non-directive and open-ended, while Prof. Paxton's seemed to be more pointed and specific.

In an effort to capture these differences analytically, we developed and applied the coding scheme described above. Due to space limitations, we provide in Table 2 only a few of the eighteen coding categories and counts for the instructors' questioning. Questions coded as *Thinking*, for example, offered invitations or openings for students to describe their thinking or share their opinion. These were classified as one of three types: *Opening*, *Neutral*, or *Narrowing*:

- *Opening*: Invites a student's opinion or thinking (or sometimes a response), but without suggesting an increasing in focus from previous comments or questions. Example: *What do you think, Stacy?*
- *Neutral*: Invites opinion/thinking, but focus is unclear from the context
- *Narrowing*: Invites a student's opinion, thinking, or response, directed toward a specific idea that has emerged from the conversation; tends to focus the conversation rather than open it up to new directions. Example: *What do you think about what Drew just said?*

Questions coded as *Math/Service* are those that pointedly ask a specific mathematical question, usually referring directly to specific mathematical objects, and usually having a right or wrong answer. These questions are posed "in service" of making mathematical progress on the task/problem at hand. (Example: *If the derivative is negative, what can we say about the function?*)

Table 1: Representative samples of each instructor's pattern of questioning and soliciting students' contributions. Prof. Gage's questioning is very non-directive and open-ended, while Prof. Paxton's is increasingly pointed and specific.

Prof. Gage	Prof. Paxton
G: Can somebody say a bit more about the variable aspect? And you may, yourself, I mean, you just said, you know…. S: (replies) G: Tom? S: (replies) G: Sue, can you explain what Tom…? S: (replies) G: Could somebody, somebody, please? S: (replies) G: To be functions? S: (replies) G: So you, Ron, what was in your mind here? S: (replies) G: I think we can go on forever with this. It's kind of interesting.	P: Do we agree with Cathy? Ss: (replies) P: So in System A when y is zero, species x and species y are getting what? S: (replies) P: So how do you go from that observation to the conclusion that in system A ... the species are cooperative? You're Kevin, is that right? S: (replies) P: So what does it mean to say that the two species are cooperating? S: (replies) P: So Keith, are you gonna jump on that? Because that's exactly the same reasoning as we heard. Conor? You don't have to raise your hands! Just talk!

Table 2: Excerpt of coding counts for instructors' questioning during whole class discussions of nearly identical lengths.

	Prof. Gage	Prof. Paxton
Questions		
Thinking	32	19
Opening	24	10
Neutral	0	1
Narrowing	8	8
Math/Service	2	26
Clarify	12	4
Progress/Assess	13	8
Justify	0	6

A comparison of the two instructors shows that Prof. Paxton's questioning was somewhat more likely to include *Narrowing* questions, and much more likely to include *Math/Service* questions. He also asked students to justify their responses at times during the conversation, while Prof. Gage never did. Prof. Gage, on the other hand, tended to keep his questions much more open-ended and non-directive, and he was more likely to ask students to clarify their ideas and comments and to *Assess* or "check in" with students to see if they were paying attention, understanding or accepting what was under discussion. A complete analysis of the whole-class discussions shows similar differences in the nature of the instructors' comments, as well. In sum, the two instructors demonstrated markedly different types of engagement and guidance during whole class discussions, contributing to what we identify as markedly different enactments of the curriculum and opportunities for student learning.

Opportunities to learn

The two instructors' different approaches to questioning and participating in the whole-class discussions offer a first measure of how, despite sharing the same written curriculum, very different enacted curricula arose in the two classrooms. In what follows, we will present additional, significant differences in the enacted curricula, but as a means of presentation, we find it helpful to present these data in the context of considering how these differences may have affected student learning. Unfortunately, we do not have the data available to directly study student learning through assessment or interviews. However, we propose that considering students' *opportunities to learn* provides a means to begin to think about the big picture of curriculum enactment as it affects student outcomes. Bennett (1987) discusses a student's *opportunity to learn* as a function of several variables, including a student's time on task, the nature and quality of the task, and the degree of student engagement. In the context of Prof. Gage's and Prof. Paxton's classes, we assume that students opportunities to learn would be affected by the time given to particular topics and discussions, the relative emphasis or value placed by each instructor on the matters discussed, and the degree to which students themselves demonstrated engagement in the class. We emphasize that the question is not simply whether students learn "more" or "less," but that as instructors influence different curriculum enactments, the nature and character of what is being learned can be affected as well. This analysis cannot replace a more direct assessment of students' actual progress, but we believe it points to the research need for RUME studies of teachers' practice to give due attention to student assessment.

Instructor questioning and commentary

Our analysis above suggests to us that the differences in the ways that Prof. Gage and Prof. Paxton participated in their whole-class discussions likely resulted in offering the students in the two class differing opportunities to learn. We can hypothesize what might be considered both "pros" and "cons" to each instructor's approach. Prof. Gage's more open-ended and less pointed questioning, for example, may have encouraged a wider variety of students' ideas to be placed in conversation, allowing students to consider more perspectives on the same problem; it might also have extended the conversation and made it more challenging for the class to reach consensus on the problem at hand. Prof. Paxton's more directive line of questioning, on the other hand, may have encouraged a common focus, permitting a quicker achievement of consensus; but it might also have discouraged alternative ideas or solution approaches from emerging, thereby favoring the instructor's preferred solution or explanation over alternatives. Our point is not to determine which, if any, of these hypotheses is correct, but to emphasize that *we really do not know* how the different instructor contributions affected the enactment of the curriculum either to the benefit or detriment of student learning.

Turns of talk

How students engage with a topic may be analyzed in terms of how often they actually participate in the whole class discussions. We understand that this may be a simplistic view of "engagement," as students who do not speak may or may not be actively engaged in classroom learning. However, research does show that participation in the mathematics discussion is an effective way for students to construct knowledge (Webb, 1991; Cobb, Stephan, McClain & Gravemeijer, 2011). In Table 3, we report the count data for turns of talk by teachers and students that occurred on the first day of each of the professor's class during all periods of whole class discussion.

Table 3: Analysis of classroom "talk" during whole-class discussion

	Prof. Gage	*Prof. Paxton*
Turns of talk	188	183
Instructor's turns of talk	74 (39.4%)	87 (47.5%)
Students' turns of talk	114 (61.6%)	96 (52.5%)
Students in class	22	21
Students contributing	16 (72.7%)	9 (42.9%)
Students making 90% of contributions	10 (45.5%)	5 (23.8%)

A "turn of talk" refers to the spoken contribution of an individual speaker, beginning after a previous speaker concludes and ending before a subsequent speaker begins. (Occasionally, overlapping contributions are coded as individual turns.) As one can see, Prof. Gage's and Prof. Paxton's classrooms had almost an identical number of turns during their classes' whole-class discussions (which lasted almost identical lengths of time). Professor Paxton made 47.5% of the utterances and Prof. Gage made 39.4%, giving Prof. Paxton a somewhat more frequent role in the discussion. This indicates that there were a few more times in Prof. Gage's classroom than in Prof. Paxton's during which the whole-class discussion continued among students without a contribution by the professor. This corresponds to the fact that, throughout the whole-class discussions, more students made vocal contributions in Prof. Gage's class (16) than in Prof. Paxton's (9). A smaller number of students in Prof. Paxton's class (5) tended to dominate the conversation by contributing 90% of students' turns of talk, whereas in Prof. Gage's class, contributions of talk were distributed more broadly among 10 students. These differences in student participation rates suggest the possibility of corresponding differences in students' opportunities to learn.

Content analysis

The two mathematics professors addressed the first day's mathematical content in different ways, and we contend that how the professors interpreted the curriculum is an important and appropriate way to consider how the enacted curriculum may have affected student learning. In terms of the opportunity to learn, this content analysis considers both students' time on task and the nature and character of the tasks with which they engaged. Table 4 presents a brief description of how each professor addressed the mathematical content provided by the written curriculum materials.

As Table 4 shows, Prof. Gage had his class take up only the first task of the curriculum on Day 1, and the class ended without the students reaching an apparent consensus on its solution. He spent significant time (over 13 minutes) dealing with the question, "What are x and y?

Functions? Variables? Numbers?" Although our focus is on Day 1, we note that Prof. Gage did not return to the first task on Day 2, but he did continue the discussion of the "What are x and y?"

Table 4: Overview of how the class was conducted and details of time on different content

Prof. Gage	Prof. Paxton
• Students spent the entire first class working on the introductory task. • A significant amount of time was spent on the question, "What are x and y?" • Students broke into groups four separate times to discuss the first task, interspersed with whole class discussions. • No stated or assumed consensus was reached on the first task.	• Students completed the first task with "announced consensus." • A very brief discussion of the question "What are x and y?" took place. • Students broke into groups once for each of three consecutive tasks. • Class completed the second task with "announced consensus." • Class worked on third task with uncertain consensus at end of class.

Time spent on content (italicized times in bold taken from Day 2)		
	Prof. Gage	**Prof. Paxton**
Preliminaries	31:10	13:27
Problem 1	45:50	29:08
Small group	10:15 (4)	13:21 (1)
Whole class	35:35	15:47
What are x and y?	13:13 ***16:30***	1:28
Problem 2	***55:26***	15:52
Problem 3	- -	16:41

question on the second day for an additional 16.5 minutes, suggesting the significance that he placed on students' engagement with that particular question. This question was listed in the teacher materials provided by the IO-DE authors (and thus was a part of the written curriculum) as an idea that is important to address with students during the early classes. Prof. Paxton had the same set of instructor materials, and he, too, raised the question about x and y, but he spent only about 1.5 minutes on the matter without returning to it on either of the first two days.

In contrast, Prof. Paxton facilitated the completion of the first task of the curriculum, with announced consensus. By "announced consensus," we mean that Prof. Paxton made a clear attempt to ask the class if everyone agreed with the conclusions that had emerged from their discussions, and, receiving no negative feedback from students, he indicated that it was time to move on to the next problem. He then asked students to work on the second task, first in small groups and then in whole-class discussion, again reaching announced consensus. He then had students work on the third task of the curriculum, and although he identified at the end of the class the correct answer to the problem posed by the third task, Prof. Paxton noted in his interview that it was not clear that the class at arrived at consensus around it.

What is clear, however, is that by the end of each instructor's first day of class, the students had been exposed to very different mathematical content, even though both were using the same curriculum materials. Prof. Gage's class focused entirely on one problem, and the question of "What are x and y?" received significant attention. Prof. Paxton's class addressed three separate tasks, and the question about x and y was addressed only very briefly.

Additionally, the difference in times devoted to small group and whole class discussions may have provided different opportunities for students to engage in learning. Professor Paxton provided students with considerably more time than Prof. Gage in small group discussion (26 minutes, spread over three tasks), using small group discussion only at the beginning each task. Professor Gage provided less time small group discussion (a bit over 10 minutes), and distributed it in smaller amounts over four separate periods during the consideration of the first task. Because we do not have records of what went on in the small groups discussions, we really cannot hypothesize how these differences may have affected student learning.

The results from this section show how different enactments resulted in very different opportunities to learn. On one hand, significantly more students were engaged in the whole class discussions in Professor Gage's class and spent more time thinking about the key question, "What are x and y?" On the other hand, Professor Paxton's students were able to consider more mathematical tasks and topics. We believe that there is strong evidence that the students' *opportunities to learn* differed significantly across the classes. We cannot, however, determine how actual student learning differed. But that is precisely our point.

Discussion

We approach the study of two undergraduate mathematics professors' practice by comparing their enactment of the curriculum and resulting opportunities for students to learn. Our concern is that RUME researchers may have become too complacent in assuming that the use of a particular written curriculum can be readily correlated with expectations about the nature, the quantity, and the quality of what students are actually learning. We believe that it is possible to overestimate the causal relationship between a written curriculum and subsequent student learning. Rather, we claim that the enactment of the curriculum, which is constituted in the classroom by both the professor and the students, is particularly appropriate to study as we look at student outcomes. We propose that greater attention needs to be given to distinguishing between the different curricula at work: the written, the intended, and the enacted curricula offer a powerful way to reconsider the connection between curriculum materials and student learning.

We presented an analysis of data drawn from the practice of two mathematics professors on the first day of an undergraduate differential equations course. Both professors were teaching for the first time using a new curriculum (for them) with a student-centered and inquiry-oriented pedagogical focus. The written materials consisted of a sequence of student tasks and some informal teacher notes about implementation from the curriculum's author. We found that, even though the classes appeared the same when considered at a coarse grain size, significant differences could be identified concerning the instructor's role in discussion, student participation, the pace and quantity of the material discussed, and the emphases placed on different aspects of the mathematical content included in the curriculum materials.

The kinds of questioning that the instructors used were different. Even though both professors asked what we call "thinking" questions, Prof. Gage used more *opening* questions and Professor Paxton used more *narrowing* questions. This created different kinds of learning spaces as the students in the two classes participated in the curriculum enactment. Additionally, the cycling of small group and whole-class discussion was different. Prof. Paxton began each task with one small group discussion and then conducted whole class discussion to complete and come to what he perceived as consensus on two curriculum tasks, with additional progress made on a third. Professor Gage used of a series of small group and whole-class cycles during which the small group sessions lasted for short periods of time. Consensus around the first task of the curriculum discussed by his students was not reached.

We hypothesized differences in students' opportunities to learn in the two classes by considering their engagement, their time devoted to the tasks, and the nature and character of the tasks and supporting discussion. The two classes demonstrated different levels of student engagement, different emphases on the mathematical content discussed, different amounts of time spent on individual topics and tasks, and different rates of progress through the curriculum materials. Although we have no direct measure of student learning in the class, it is difficult to image that the two instructors' use of the same curriculum materials resulted in comparable student experiences. Furthermore, we do not think it helpful to judge that either of the two instructors succeeded or failed in making good or "faithful" use of the curriculum materials. We believe that differences like those we observed are to be expected to varying degrees when any individual instructor engages in transforming a written curriculum into an enacted one.

Conclusion

The RUME community has produced a wealth of research in recent years on the experiences of instructors and students in what is often called "research-based" curricula. Very often, however, the fact that these curricula emphasize inquiry, investigation, and other student-centered activities is taken as sufficient for presuming their ultimate benefit to students. We do not doubt the consistent research findings that student engagement and inquiry are beneficial. However, we are concerned that a large amount of the research on instructors' use of such curricula either does not attempt to link the instructors' activities to student outcomes, or appears to presume that "faithfulness" or "fidelity" to a curriculum as intended by its authors is possible and to be expected. We fear that as a research community we may be too easily taking student outcomes for granted, rather than taking on the admittedly daunting task of asking harder and deeper questions about how the myriad ways that any written curriculum may be enacted ultimately impact our students' experiences. It is necessary to promote research that directly connects curricula, teacher practice, and student learning, rather than isolates them.

References

Ball, D. & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education, 60*(5), 497-511.

Bennett, N. (1987). The search for the effective teacher. In Delamont, S. (Ed.), *The primary school teacher.* Lewes, Falmer Press, 51-56.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2011). Participating in classroom mathematical practices. In *A Journey in Mathematics Education Research* (pp. 117-163). Springer Netherlands.

Department of Commerce (2012). The competitiveness and innovative capacity of the United States. Report retrieved 08/08/2013 at http://www.commerce.gov/americacompetes.

Fukawa-Connelly, T. (2012). A case study of one instructor's lecture-based teaching of proof in abstract algebra: Making sense of her pedagogical moves. *Educational Studies in Mathematics, 81*(3), 325-345.

Johnson, E. Caughman, J. Fredericks,, J. & Gibson, L. (2013). Implementing inquiry-oriented curriculum: From the mathematicians' perspective. *Journal of Mathematical Behavior, 32*(4), 743-760.

Lee, H. S., Keene, K. A., Lee, J. T., Holstein, K., Early, M.E., Eley, P. (2009). Pedagogical content moves in an inquiry-oriented differential equations class: Purposeful decisions to further mathematical discourse. *Proceedings of the 12th Annual Conference on Research in Undergraduate Mathematics Education conference*, Raleigh, NC.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33-84.

Patterson , J.,Thomas, M. & Taylor, S. (2011): Decisions, decisions, decisions: what determines the path taken in lectures? *International Journal of Mathematical Education in Science and Technology, 42*(7), 985-995.

Pinto, A. (2013). Revisiting university mathematics teaching: A tale of two instructors. Presented at the Eighth Congress of European Research in Mathematics Education (CERME 8), Antalya, Turkey.

Rasmussen, C. (2006). Unpublished curricular materials for differential equations.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, *75*(2), 211-246.

Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientations toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Education, 35*(5), 352–388.

Ridgway, J., Zawojewski, Z., Hoover, M., & Lambdin, D. (2003). Student attainment in the connected mathematics curriculum. In *Standards-based school mathematics curricula : what are they ? what do students learn ?* New Jersey: Lawrence Erlbaum, pp. 193-224.

Speer, N. M., & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. *Journal for Research in Mathematics Education*, *40*(5), 530-562.

Speer, N. M., Smith, J. P., III, & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, *29*(2), 99–114.

Stein, M., Remillard, J., & Smith, M. (2007). How curriculum influences student learning. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 1, pp. 319–369). Charlotte, NC: Information Age Publishing..

Tarr, J. E., Reys, R. E., Reys, B. J., Chávez, Ó., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education, 39*(3), 247-280.

Trenholm, S., Alcock,L. & Robinson, C. (2012) Mathematics lecturing in the digital age, *International Journal of Mathematical Education in Science and Technology*, 43:6, 703-716.

Wagner, J. F. (2007). A comparison of two mathematicians' use of an inquiry-oriented differential equations curriculum. Presented at the Tenth Conference on Research in Undergraduate Mathematics Education (RUME), San Diego, CA.

Wagner, J. F., Speer, N. M., & Rossa, B. (2007). Beyond mathematical content knowledge: A mathematician's knowledge needed for teaching an inquiry-oriented differential equations course. *The Journal of Mathematical Behavior*, *26*(3), 247-266.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*(5), 366-389.

Wells, G. & Arauz, R. M. (2006). Dialogue in the classroom. *Journal of Learning Sciences*, *15*(3), 379-428.

STUDENT CONCEPTIONS OF TRIGONOMETRIC IDENTITIES

Benjamin M. Wescoatt
Valdosta State University

Although the research literature concerning topics in trigonometry is growing, explorations of students' conceptions of trigonometric identities is scant. This study aimed to contribute to this area by considering the extent that students developed a structural understanding of trigonometric identities and the implications their understanding had on the solving of problems involving identities. Through task-based interviews involving verifying trigonometric identities, students appeared to view certain identities differently than other identities, holding a deeper understanding of them. Their understanding and trust in these identities allowed for the students to use them in the construction of new identities. In order to verifying identities involving complicated function arguments, some students either ignored the argument or replaced it with a single letter, typically an x. Although doing so allowed the students to focus on the structure of the identity, students' conceptions of variable remained at a conceptually-weak level.

Key words: Trigonometric identity, Variable, Function argument, Procedural-structural

Although research into concepts in trigonometry has historically been sparse in comparison to other areas of mathematics, over the past decade, the body of research has slowly accumulated. In her dissertation, Brown (2005) explored the development of a framework for necessary components in students' understanding of the sine and cosine function. She believed that for students to have robust understandings, they needed to have fluid, integrated views of the functions as representing directed distances, y-coordinates, and triangle ratios. These views were all grounded in the context of rotations in the standard position. Additionally, students should have a connected understanding of the functions as represented by calculator outputs, right triangles, rotations, and sinusoidal graphs.

Taking a break from his research into proof, Weber (2005) investigated how students experienced the sine function in a prescriptive sense. Trigonometric functions are different from initial types of functions, such as polynomial functions, that students encounter in that they do not have an explicit form. This issue can become a cognitive barrier in that students initially conceive of functions as procedures (Breidenbach, Dubinsky, Hawks, & Nichols, 1992). That is, functions have a prescriptive nature; the student inputs a value, the function outputs a value. Weber described how students could approach trigonometric functions in a prescriptive sense through the measurement of triangle sides and the construction of ratios. The process allowed these students to gain a deeper understanding of the trigonometric functions and use their understanding of the process to justify properties of the functions in comparison to students who did not undergo the treatment.

Moore (2014) described the importance a robust understanding of angle measure plays in a student's understanding of the trigonometric functions. Within the context of co-variational reasoning, the student in the study, Zac, was able to construct a process image of the sine function. Zac reasoned about the circular motion inherent in the unit circle and connected it with the dynamic nature of the sinusoidal shape, allowing him to recreate the graph without relying on a memorized image or numerical calculations.

As the previous research shows, most of the research concerning students' understanding in trigonometry has been directed toward the trigonometric functions. A next step would be to explore students' conceptions of trigonometric identities. A trigonometric identity is an object typically encountered by many high school or college mathematics students. The identity itself is a tautological statement claiming that two expressions composed of certain combinations of trigonometric functions actually describe the same underlying mathematical object despite appearing to be different. The notion of trigonometric identity is rich in mathematical conceptions. Students should be able to coordinate ideas of function and equivalence in order to understand the intended meanings imparted by the identity. Additionally, students should have a mature conception of the function argument, understanding the equivalence of the expressions for all inputs in the domain of the expressions. Finally, if students engage in the activity of verifying identities, students must understand what constitutes valid mathematical argumentation.

The studies of Weber (2005) and Moore (2014) focused on students developing a process conception of a trigonometric function. As students explore identities, they need to move beyond this conception view and focus more on the structural aspects of the functions involved. This study aims to fill in the gap in the literature pertaining to students' conceptual understanding of trigonometric identities. In particular, this study explored if students do begin to view trigonometric identities as indicating structural relationships rather than viewing identities in a procedural sense and examined the implications this shift could have for students in terms of solving associated problems.

Theoretical Framework

Rooted in the theories of Piaget, APOS theory attempts to describe how students may come to understand certain mathematical objects. Underpinning APOS is the hypothesis that

> an individual's mathematical knowledge is her or his tendency to respond to perceived mathematical problem situations and their solutions by reflecting on them in a social context and constructing or reconstructing mathematical actions, processes and objects and organizing these in schemas to use in dealing with the situations. (Dubinsky, 2000, p. 11)

According to Dubinsky and McDonald (2001), an action is a learner-perceived external transformation of an object. Actions usually occur in a step-by-step fashion with reliance on memorized procedures. Once the learner has repeated an action, reflection on the action may interiorize the action into a process. A process does not need to be physically performed; the learner may envision the process and the result. Thus, a process does not have a reliance on external stimuli but is under the control of the learner. Once the learner understands that a process represents a totality, the learner is said to have encapsulated the process into an object. As an object, the focus has shifted from a procedural understanding to a structural understanding of the concept. The learner constructs the mathematical concept's schema by collecting into a coherent framework all of the other actions, processes, objects, and schemas associated with that concept. APOS theory was used to understand trigonometric identities as moving from a process conception to an object/structural conception.

Küchemann (1978) describe the many conceptions of numerical variables that students held when solving problems. He described these uses as being a letter that was evaluated, ignored, or treated as an object. Additionally, he classified them as letters treated as specific unknowns, generalized numbers, or actual varying quantities. Gray, Loud, and Sokolowski (2009) characterized the first three conceptions of variable as arithmetic or procedural thinking. In other words, students could solve problems by implementing a particular process. The later three conceptions were characterized as algebraic thinking; students using those interpretations focused on the structural aspects of the problem. In this study, uses of literal symbols, or a variable, will be characterized as being either arithmetic or algebraic. That is, students' conceptions of variable will be interpreted as either procedural or structural, with structural thinking considered as representing a higher cognitive level.

Methodology

The data for this study were collected from a college trigonometry course at a large research university as part of a larger case study on verifying trigonometric identities. Thirty-three students participated in the study, responding to prompts involving verifying identities and solving verification problems on homework, quizzes, and exams. Of these thirty-three students, eight agreed to participate in individual task-based interviews. Each interviewee solved problems while speaking aloud his or her thought processes. The pseudonyms of the participants in the interviews were Alan, Amber, Bella, Charles, Cooper, Helen, Katie, and Maria. Charles could be classified as deficient in his mathematical understanding and ability; the remaining interview participants were either above average or excellent in their understandings and abilities. The audio from the interviews was captured and subsequently transcribed. Additionally, student work generated during the interviews was retained for alignment with the transcripts.

During the transcription process and while reading each interview, general themes were jotted down in an open-coding process; the approach could be described as being in the style of Grounded Theory. Each evolving theme was compared to themes from previously analyzed interviews. The themes were triangulated within and across interviews. After the initial analysis of each interview, the interviews were further analyzed and the themes modified until the themes adequately described the observations from the interviews.

In pure Grounded Theory, no theoretical framework is used to initially analyze the data. Instead, the data construct the framework. For the purposes of the data analysis for this study, initially no framework was used. However, as the themes began to evolve, it became apparent that pre-existing theory regarding the shift from procedural to structural understanding would be useful in framing the results. Thus, the procedural-structural framework was utilized to interpret the themes.

Results and Discussion

In this section, results of the task-based interview will be shared and discussed. Again, the focus of the analysis was on how students may understand the trigonometric function as an object.

Encapsulation of the Trigonometric Identity

Students appeared to encapsulate a core of well-known identities. These identities became well-understood by the students through repetitive successful utilization. Students were quite ready and comfortable to use these core identities to solve problems. Furthermore, some of the students apparently conceived of these identities differently than other identities.

An example of how students perceived of these core identities differently is best understood by examining comments made by Cooper. The first question asked in the interview was the following: "Do you consider the following equation to be an identity?"

$$\tan^2 x = \frac{\sin^2 x}{\cos^2 x}$$

While students may have believed the equation to be or not be an identity for various beliefs, Cooper had an interesting and enlightening response.

> COOPER: This, I, I would consider this to be a definition, I think. Cuz tangent x, I mean, tangent quantity squared, tangent is sine over cosine. I consider that a definition.

Not only did Cooper not believe the equation to be an identity, he considered it as representing some other class of object, an object for which he had named, a *definition*. Through prompting, Cooper clarified what a definition meant to him.

> COOPER: It's kind of like looking at a word. I mean, you've got different words have different meanings. And if you look at tangent, I, I look at it as like, you know, you know, something that describes something. … Like, you know, whenever you think of just definition of a horse, or something, it's just an explanation of what it is. So whenever you think of tangent, you're kind of explaining what it is, well that's sine over cosine.

Thus, for Cooper, definitions focus more on underlying aspects, referring to alternative ways to describe a particular object. In this example, he can describe an object as the tangent function or as a ratio of the sine and cosine function.

Cooper was able to articulate differences between how he viewed definitions and identities. Thus, in his mind, a definition and an identity were distinct objects. In exploring what he considered to be identities or definitions, Cooper claimed that identities differed from definitions in how he recalled them to consciousness.

> COOPER: They're [identities] not something that kind of just rolls off the thought. You know, its' not something that instantly pops up in your head. … Definitions are automatically thought of. Identities take me a moment to catch onto.

For Cooper, the hesitation in invoking an identity appeared to originate in a need to perform a manipulation of sorts. When distinguishing something as an identity, he mentioned process that occurred, usually mentally. As an example, while he considered $\sin^2 x + \cos^2 x = 1$ to be a definition, defining what 1 was, he viewed $\sin^2 x = 1 - \cos^2 x$ to be an identity.

> COOPER: It doesn't really just roll off the head. And it's not, it's kind of, there's still manipulation of your original Pythagorean identity. … I look at the original problem and I'm imagining, you know, going ahead and changing it up, and, so, it's, I usually do that so I don't get it wrong and like kind of make a mistake somewhat.

Thus, if Cooper needed to employ a process of change in order to arrive at an equation to use, he considered this equation to be an identity. He knew that the equation was true as he began with what he considered to be a definition and manipulated that. In a sense then, Cooper knew the definitions to be true. He believed the descriptions they contained to be true descriptions. Therefore, he could use these truths, the definitions, to arrive at other truths, the identities; he could create new truths from the old, known truths. Hence, Cooper was quite comfortable with what he considered to be definitions.

Conceiving of an equation as a definition allowed for Cooper to proceed in a problem without worrying about the correctness of the step. He did not hesitate. The distinction between definition and identity was important for him.

> COOPER: It's easier to think of it as a definition because I just whenever I see something I like to go look at it and immediately know that it's something else. Like, the definition just, whenever you memorize it like that, I think it makes it a little more easier. … You simplify things like that, I think just looking as a definition, kind of helps out a bit.

Thus, Cooper valued this distinction he held. He was provided a measure of confidence knowing the definitions.

The development of definitions in Cooper appeared to originate in the repeated utilization of certain identities.

> COOPER: The Pythagorean identities is the only ones I can really memorize for the most part. And those just are the most, their usually the ones that I use the most. Uh, the double angle identities, those are sometimes useful, but I haven't really used them a whole lot. And the angle sum, haven't really used those too much. And the angle differences, those I always forget. … The Pythagorean identities. Yeah, those are the easiest ones to remember. Those are the most stuck on because you use them all, I, I

Table 1
Cooper's Self-Identified Definitions and Identities in the Interview

Classification Name	Equation
Definition	$\sin^2 x + \cos^2 x = 1$ $\tan x = \dfrac{\sin x}{\cos x} \quad \cot x = \dfrac{\cos x}{\sin x}$ $\tan^2 x = \dfrac{\sin^2 x}{\cos^2 x} \quad \cot^2 x = \dfrac{\cos^2 x}{\sin^2 x}$
Identity	$\sin^2 x = 1 - \cos^2 x$ $\tan x = \dfrac{1}{\cot x}$ $\csc^2 \theta = \dfrac{1}{\sin^2 \theta}$

> think you use them a lot more than you do the other ones. … You know, rinse and repeat, practice makes, you know, you memorize things if you keep using them.

Thus through it all, Cooper referred to using particular identities over and over. These were the ones that he *really* had memorized. Table 1 includes equations that he distinguished as being definition or identity during the course of the interview.

While not specifically providing names, other students indicated a view similar to Cooper's view regarding well-known identities. In discussing what a trigonometric identity meant to her, Amber spoke of having a fluid understanding of them.

> AMBER: Once I see the connections, things just flow. So, I mean, once you've, like, once I know the identities, like, inside and out, or basically know them without having to think about them, like, oh, wait a minute, what this equal to, or whatever?

For Amber, intimately knowing an identity allowed her to readily use this identity. As Cooper experienced, this intimate understanding developed through repeated usage.

> AMBER: I didn't see an identity straight off that I could use besides the reciprocal identity. Um, of course I don't even think of that as an identity. But, um.
>
> INTERVIEWER: Why not?
>
> AMBER: It's just second nature from high school. … The reciprocal identities are so simple that I just, it, they're just.
>
> INTERVIEWER: You just do it.
>
> AMBER: Yes. It's one of those things that I don't, I have to sit there, oh it's an identity. When I'm like, if when I'm talking about it, it's like, oh, it is an identity. But, um, it's just like something I just know and kind of go with it.

Interestingly, Amber forgot that well-known identities were in fact identities. In a way, she no longer viewed these identities as identities. She was able to use them without consciously thinking about their validity because she already knew they were true. She had repeatedly used them, successfully, since high school.

Having confidence in the identity was important in problem-solving effort for Amber. Even if she verified that a particular equation was in fact an identity, she remained hesitant to use it in solving a problem or verifying another identity.

> AMBER: I'm always using like one of the more basic identities to make the move. I'm not using an identity I've verified. So um, I'm always using the ones that are given essentially like in the book, or something to make the move. So I'm not using identities that I've proven to, uh, verify another identity.

Perhaps Amber's qualms originate from the lack of confidence in this identity. After all, unlike the identities given in the book that she has used over and over again successfully to verify identities, she has not used this new identity yet in a verification situation.

During her interview, Katie also expressed notions of identity related to Cooper's definition conception, using language similar to Cooper's. Katie described her thought process in claiming that $\tan^2 x = \sin^2 x \,/\, \cos^2 x$ was an identity.

> KATIE: I mean I feel like I always know as soon as I look at tangent that it's sine x over cosine x.

Katie drew an immediate connection between the tangent function and the sine and cosine functions. She speculated about the origins of this immediacy.

> KATIE: I'm more used to using sine and cosine, and it's more common to me. And so I think it's the first thing that comes to my head.

Thus, as Cooper and Amber described, Katie believed repeated usage of the particular identity developed the association between the functions. These immediate connections existed for other identities for Katie. The second question asked in the interview was the following:

> Verify that:

$$\frac{1}{1-\cos^2\theta} = \csc^2\theta$$

Katie discussed why she immediately began by substituting in the expression $\sin^2\theta$ for $1-\cos^2\theta$.

> KATIE: That is like a first identity that pops into my head whenever I see one minus cosine squared that, that it's the same thing as sine squared theta.

For Katie, well-known identities *popped* into her head the same way definitions *popped* into Cooper's head.

For Katie, confidence in using certain identities seemed to relate to whether or not they existed in her core of identities. For example, the fifth interview prompt asked the following: "What does the following expression equal?"

$$\frac{\sin 2x}{\sin x}$$

In discussing her solution, Katie admitted her initial discomfort.

> KATIE: Sine two x is kind of more foreign to me than like sine squared x. … I feel like we didn't do as much work with those, with, like, a double angle identity as I did a Pythagathorean [sic] identity. So I was, I don't know, I just. I don't know those off the top of my head.

Katie related comfort in using identities with how readily she recalled the identities. The lack of using identities proved a barrier for her. She described why she believed the problem to be tricky for her.

> KATIE: It was using a double angle identity, which I'm not, it's, uh, more, um, foreign to me.
>
> INTERVIEWER: Okay. So do you feel then knowing, or being more familiar, having memorized identities makes it easier for you to solve problems?
>
> KATIE: I mean, I guess it just makes me feel more comfortable when I am working it instead of having, I guess, to go look at a chart and figure out what it is and then coming back. … I guess you just feel more comfortable with it. When you have it in your head, that means you've probably worked with it many, many times. … I mean, like I said, the cosine and sine, we used those from the start of identities to the end of identities. And you used them repeatedly.

In addition to describing certain identities as *popping* into her head, Katie also expressed the view that she forgot certain well-known identities were in fact identities.

> KATIE: Change like, say like cosecant x to one over sine x. Or like change it, I guess that is an identity. But I consider that just saying they're the same two things, just right there. You know, you can use this one or this one, whichever one you need to use at the time.

Thus, Katie described this well-known identity as describing a relational aspect. She knew the two expressions described the same underlying concept. She was empowered to use either expression, depending on which one was beneficial for the situation.

Cooper and Katie alluded to difficulties in employing certain identities that were lesser-known to them through lack of usage. Specifically, they mentioned the double-angle identities. Lack of knowledge of identities presented barriers for some students. Helen described such a situation as she attempted to verify the identity $\sin 2y = \tan y\,(1 + \cos 2y)$.

> HELEN: I still don't fully understand what cosine of two y or sine of two y means. That's so foreign to me that I would want to avoid it. But I can't in this problem. Like it's unavoidable. So, I don't understand the problem.
>
> INTERVIEWER: So you feel not fully understanding what cosine two y means is really making it hard for you to work this problem?
>
> HELEN: Yeah, or sine two y. … Like it confused me. Even though like if I was to eventually understand what that means, then I could work it easily. But, if I'm hung up on something at the very beginning, it's really hard to move past that. Because, this whole problem depends on my understanding of the problem itself.

For Helen, she was unable to overcome her lack of understanding of the identity that she needed to use. This lack of intimate knowledge inhibited her progress on the problem. Without having the conviction about the necessary identity, she was stuck.

Having intimate knowledge of identities, possessing a core of well-known identities, was important for students' problem-solving efforts. Like Cooper, many students expressed the idea that repetitive successful use made these identities easier to use. Students were able to manipulate these *definitions* at will in order to best suit the problem. Thus, students were able to generate new identities from the old, well-known identities, focusing on the structural aspect of the identity.

An inability to generate these new identities, deviating from the known identities, was a hallmark of a student who struggled during the interview, Charles. Charles repeatedly stated that he did not have the identities memorized, indicating a lack of intimate knowledge of them. Additionally, Charles could not manipulate common identities provided to interview participants in a table. The following exchange comes from his attempt at solving the second prompt. At this point in the interview, he was struggling with what step to take.

> INTERVIEWER: You're saying one minus cosine squared theta equals sine squared theta.
>
> CHARLES: Uh.
>
> INTERVIEWER: Do you see that identity, or one similar to that identity?
>
> CHARLES: No. One similar?
>
> INTERVIEWER: Mm-hm. Maybe it's not exact, but.
>
> CHARLES: Um. I mean there's the one minus two sine squared. And the, I mean nothing really stand out to me.

Charles attempted to visually match the expression in the problem with one on the identity sheet. He was unable to manipulate the Pythagorean identity, despite it being pointed out to him.

CHARLES: It's not exact, but I don't know. It doesn't, like, I wouldn't have thought to have used it in this.

For Charles, the lack of repeatedly using identities and developing a sense of intimate knowledge about them hindered his ability to use them and manipulate them. Through a discussion of the aftermath of the problem, Charles reiterated his inflexibility with the identities.

CHARLES: If I had something to go off of, I would feel a hundred percent better in writing it down. Like if I had a list that, you know if I had one of those cards, and mine would have one minus cosine equals sine [sic]. … I wouldn't have thought to like looked at it that way, I guess. When, Like, I guess when I look at something, I'm like, well, that's that. And, uh, a lot of the time, I forget to do things, like, oh, I can put, or, I, I wouldn't see that you could subtract cosine from each side.

Unfortunately throughout the interview, Charles repeated the pattern of attempting to find an exact match for the expression on the identity sheet, displaying an inability to create the needed identity from existing identities.

The Function Argument

In treating identities as objects, attending to the structure and making use of the equivalence of expressions while verifying identities, students could ignore the process suggested by the identity. In particular, students could generally ignore the inputs of the functions involved. After all, identities were true for all input values in the domain of the functions. As a result, students treated the function argument in interesting but perhaps debilitating ways.

For the fourth interview prompt, students were asked to verify that the following equation was an identity:

$$\frac{1}{1-\cos^2(2\alpha-1)} = \csc^2(2\alpha-1)$$

Students handled this problem in similar ways. Maria described how she mentally viewed the function argument.

MARIA: What is in the parentheses is just one term.

INTERVIEWER: Okay. So wait, how do you see that even though visually it's like two alpha minus one. How do you see that in your head? Do you see a two alpha minus one?

MARIA: It's x.

INTERVIEWER: Do you, so in your head, do you, do you see an x?

MARIA: So, everything in the parentheses is just an x.

INTERVIEWER: So do you actually in your head. You're thinking of x?

MARIA: I really do.

In order to verify the identity, Maria needs to envision the function argument as nothing more than an x. She claimed that doing so simplified the problem for her.

While some students used a similar approach to Maria, actually writing $\cos^2(x)$, other students completely ignored the argument. Helen commented on her reasoning for using exactly this strategy.

HELEN: I'm not mentioning it because I didn't really do anything with it. … Ignore it just because it can get me a lot more mixed up than I need to be. … I realized that nothing has anything to do with it. … I kind of plugged it in at the last minute. … But it didn't have any meaning as to I didn't have to foil it out or anything. So, just keeping it as like something that's like x.

In Helen's view, since she did nothing with the argument, she was justified in ignoring it. By ignoring the argument, something that was a little complicated for her, she was able to

$$\frac{1}{\boxed{1-\cos^2}(2\alpha-1)}$$

Figure 1: Cooper's focal point.

negotiate, in her opinion, the verification of the identity. She merely needed to reinsert the argument, tacking it on, at the end. Cooper used an extreme method to ignore the complicated function argument; he explicitly drew a box around what he needed to focus on for solving the problem, omitting the function argument (Figure 1).

> COOPER: I kind of took that one out of mind. … I realized that both sides had the, had the same thing and that it really wasn't any much of an effect of the actual, original, what you're trying to verify.

As was the case with Helen, Cooper felt justified in ignoring the function argument. Later in the interview, he clarified his stance on the function argument.

> COOPER: I like to just ignore that one unless there's multiple variables in the situation … it really doesn't play a whole lot of importance unless there's multiple variables in it. And as long as you put it in your final answer … it's still right.

Again, his comment echoed Helen's opinion of the function argument. As long as the argument was tacked back in the problem, the verification would be valid. This cavalier attitude concerning the argument was further explained by Amber.

> AMBER: You get so caught up in the sines and cosines that sometimes the x just kind of disappears in your head. … The x is basically saying there's a variable there. ... You can't have a trig function without a variable. … All that letter is standing for is, it's standing for some number.

Thus, in focusing on what students believed mattered for the problem, the structural relationships shown by the functions, the argument became expendable; it was merely just a formality, just a technical part of the expression. After all, as Amber further explained, the verification of the identity does not depend on the symbol being used in the argument.

> AMBER: You're not actually verifying that tangent y plus cotangent y over cosecant y times secant y, um, is equal to one because of y. You're not thinking it's because of that variable. You're thinking of, in terms of the trig functions.

Hence, minimizing the function argument allowed for students to focus on the structural aspects of the expressions.

However, Amber's comments also bring to question the conception of literal symbols that students held. That is, how do students view the "x" in the function argument? Initially, Amber described the x as a generic placeholder, standing for some variable. At face value, Amber appeared to claim the x was not the variable; instead, it stood for some other variable. But then she stated the x stood for a number. Thus, rather than viewing the "x" as representing a general number, Amber seemed to view it as representing either a specific number or some other variable. Perhaps she viewed the "x" as standing in for a specific letter.

In ignoring the argument, Cooper simplified the problem for himself, allowing his focus to reside on the function. Furthermore, he discussed his view of the indeterminate nature of the argument.

> COOPER: You can always replace it with x if you want to. … It's just not to really focus on that. Cuz the main focus of the problem is what's next to it, like trying to get that idea.
>
> INTERVIEWER: Okay. So what do you mean by replace it with x?

COOPER: Uh, well, I mean, it's just kind of a placeholder. You can replace it with x, y, z, q, p, if you want. It's just a placeholder, so. That's, that's how I always looked at theta or anything. That's, uh, just a placeholder unless it's actually a constant. And then, then they have meaning.

Cooper appeared to associate the letter x as some generic placeholder, having little to no meaning for the verification problem. In his mind, what mattered was what was next to the argument. The argument had a capricious nature with the ability to be a whole litany of symbols. Interestingly, the symbol only had meaning if it actually stood for a particular value. Thus, Cooper appeared to not recognize that the symbol should have actually stood for a wide range of values, that in fact it represented a general number.

While students recognized that in fact the symbol being used in the function argument did not matter, they showed a clear preference for one particular letter.

MARIA: The first thing that comes to my mind is x. … x is like my universal variable.

AMBER: I'm like automatically kind of geared towards, oh, let's just use x.

The preference for x as the letter of choice had its roots in the introduction of the notions of symbolic algebra.

COOPER: We've always used x. … It's fourth, fifth grade, you know. Everything's just kind of always been in terms of x. … That's just usually what you always find. You know, the original first problems are like, you know, x plus three equals two.

This perceived tendency to use x continued into high school.

BELLA: Throughout my high school career, it's always been x something equals this, this, and this.

While attempting to highlight the subjectivity in choosing a symbol, teachers appeared to reinforce the notion of x as *the* variable.

AMBER: In high school, my math teacher that I actually had for eighth, ninth, and tenth grade, he told us like if we saw a variable in the book that wasn't x, just, you could change to x if you wanted to. He said he didn't care. Because it would be essentially the same thing. Because it's just a variable. Unless you have it stipulated that the variable was, was equal to something.

Again, Amber displayed a questionable conception of the symbol x, first calling it a variable, then stating that the variable was not a variable if it equaled something. She seemed to imply at this point, she lost the freedom to use a different symbol or letter.

While having the ability to recognize that the choice of symbol used was merely a choice and did not matter, students appeared to rely on using x as a crutch. That is, when expressions were presented using symbols that were not x, they felt compelled to change the symbol to an x in order to solve the problem. As previously commented, viewing the argument as an x allowed the students to shift their focus to the structure inherent in the equivalence, making the problem easier. Furthermore, students were more comfortable using the symbol x and exhibited emotion in describing their attachment to x.

MARIA: For some reason, they're easier to deal with.

COOPER: I'm just going to use x cuz that's, that works. It's going crazy. … Trying to remember it all in terms of x kind of makes it simpler.

AMBER: Dangit! Why can't you use x?

For some students, needing to think in terms of x proved a minor barrier to overcome when x was not used as a symbol in the function argument.

KATIE: I know these identities, I know them as x. And, like that's just how it's in my head, is x. So then when I see this, I kind of have to realize that that rho is the same thing as an x.

For Katie, she needed to consciously remind herself that the symbol used had no bearing on the identity. Cooper had a slightly more debilitating reaction to the lack of x.

> COOPER: When I first looked at that, and I saw that cosine two y, it didn't pop up in my mind because I was thinking of two x. … I don't see x, I kind of ignore the identities.

Both Katie and Cooper indicated knowing the identities in terms of x. That is, the symbol x acted as a placeholder for the eventual function argument in the problem. However, both of them appeared to hold onto the notion of x as being the function argument a bit too tightly, unwilling to easily replace it with the appropriate symbols in the current problem.

In order to help them negotiate solving the verification problems, student generally viewed the function argument as an unnecessary component of the problem. Accordingly, they would either drop it completely or replace it with a letter of choice, typically x. Doing so allowed them to focus on what they felt was important for the problem, the structure represented by the functions. Thus, students displayed a cavalier attitude concerning the argument, viewing it as something that was just there to satisfy a technical requirement. For some students, the symbol x took on the meaning of a general placeholder, holding a spot for the actual function argument in the problem. While students generally recognized that the truth of the identities did not depend on the symbol used, students such as Amber and Cooper appeared to display weak notions of the literal symbol, not fully viewing it as representing a general number. In fact, students could successfully verify the identity by completely ignoring the argument, considered weak conceptual understanding.

Not all students in the interviews displayed a weak understanding. Bella, assessed by the researcher to be the strongest student, was able to successfully work problems without either ignoring function arguments or replacing them with a preferred letter. Furthermore, she described a conceptually-correct notion of the role the literal symbols play in identities.

> BELLA: The x is just some number. That, it could be any number. You don't really have to know what it is to understand how the problem works.

For her, the x was a general number; it just represented some number. But her response also illustrated how the verification of identities did not lend itself to further students' conceptions of the function argument and the symbols used. Students could ignore them and still "verify" the identity.

Conclusion

Overall, students in the interview appeared to view identities in terms of their structure. They viewed a certain core of identities as representing a set of equation from which to generate new identities. Furthermore, they appeared to view these identities differently than other, lesser-known identities. These identities were easier to recall and better understood, allowing the students to confidently use them. In a sense, these students appeared to have encapsulated these identities as their core identities to use in order to verify other identities. They understood these identities on a structural level and could use them as objects.

A particular strategy some students used to focus on the structure of the identity rather than the process nature of the identity was to direct their attention away from the argument of the function. The students accomplished this feat either through completely ignoring the argument or by mentally or physically substituting a preferred letter, usually x, in for the function argument, treating the function argument as an object. The x appeared to represent or refer to something that did not matter for the problem. Thus doing so allowed the students to not worry about the argument; they could then append the argument back on the function name at the conclusion of the problem or replace the x with the original function argument.

However, in essentially ignoring the variable, students displayed a weak conceptual understanding of variable (Gray, Loud, & Sokolowski, 2009).

In the larger case study, students approached identities with complicated function arguments in a similar manner. On the exam following the unit on identities, students were asked the following question:

> Is the following equation an identity? Explain your reasoning.
>
> $$\sin^2(x^3-3) = 1-\cos^2(x^3-3)$$

A clear majority of the student (29 of 33) agreed that the equation represented an identity using the reasoning that the equation was just the Pythagorean identity. A typical explanation for the reasoning was the following, displaying the need to view the argument in terms of x to match the argument to the memorized version of the identity:

> "Because it is a Pythag. identity, which states that $\sin^2(x)+\cos^2(x)=1$ and $\sin^2(x)=1-\cos^2(x)$. It does not matter what the stuff in the parentheses or 'x' is as long as it is the same."

Some students explicitly identified the argument as being replaced by another symbol, such as the student who wrote "$x^3-1=y$" and then writing "$\sin^2 y = 1-\cos^2 y$." Other students explicitly ignored the argument, for example, writing,

> "Because the Pythagorean identity, which is $\sin^2+\cos^2=1$. Using this $1-\cos^2$ is the same as $\sin^2$ so they can be used interchangeably."

While this strategy of ignoring the given argument allowed students to match the given equation to the known Pythagorean identity, the inability to think in terms of the argument may have represented cognitive weakness. By placing the emphasis on the structural aspect of the identity as represented by the functions, students missed when a process conception of the identity and the functions would have been more appropriate to solve the problem. On a quiz given during the unit on identities, students were asked the following question:

> "Is $\cos(x)=1-\sin(x)$ an identity? Please explain your response."

Of the 25 out of 30 students indicating that this equation was not an identity, only 4 students justified their responses by choosing a particular number, substituting it into both expressions, and demonstrating that the equality did not hold for that particular value. These students were able to revert to a process understanding of identity. Of the remaining 21 students, 17 students concluded that the equation was not an identity specifically because it did not match the Pythagorean identity. That is, the equation lacked the squares on the sine and cosine function and so could not be an identity. A typical response was the following comment:

> "Because the correct Pythagorean identity is $\cos^2 x+\sin^2 x=1$, which can turn into $\cos^2 x = 1-\sin^2 x$. Since the equation given is not squared, it is invalid and not an identity."

Perhaps due to the inattention to the function argument throughout the majority of verification problems, students continued to ignore the function argument exactly when using it could have correctly solved the problem for them.

If anything, the results of this research show that while students appear to view trigonometric identities as objects, they are ill-formed objects as they could not de-encapsulate them properly; students were generally unable to revert to a process conception when needed. Furthermore, students appear to use weak notions of literal symbols while verifying identities. As Gray, Loud, and Sokolowski (2009) explored in their study of calculus students' conceptions of variables, namely, that many calculus students utilize weak notions of symbols in order to solve algebra problems, this result should come as no surprise. Unfortunately, that verifying identities could not contribute to a deeper understanding of variable is somewhat unfortunate. Gray, Loud, and Sokolowski (ibid) also found a relationship between weak notion of variable and poor performance in the course.

As a result, ways to deepen students' understandings of literal symbols used in mathematics need to be developed and implemented in the classroom. For the trigonometry course from which the students were drawn, the instructor purposefully used different symbols and letters to attempt to acclimate students to the idea of identities not depending upon the symbol of choice. Based upon the results of the interviews, students still appeared somewhat entrenched in their weak views of symbols. Phillipp (1992) suggested holding explicit discussions on the roles that symbols played in expressions. Perhaps these discussions need to occur frequently throughout mathematics classes. The goal would be to move students toward comfortably using whatever symbol or argument is within the problem. An offshoot of this issue would be the use of methods such as u-substitution in calculus. How do students view the symbol "u" during the process of solving the integral? Can students move past using a "u" and view the function arguments as objects without actually replacing them? Is this movement necessary for a deep understanding of the integration process?

One area needing further exploration is to determine the extent that the well-known core identities control the problem. For example, some students indicated a weakness when needing to implement the sum or difference identities. The question then becomes what would control the problem, students' core identities or their weak understanding of certain under-utilized items. As an example, if students were presented the problem, "Is

$$\sin^2(x-y)+\cos^2(x-y)=1$$

an identity," would they identify the Pythagorean identity structure or would they be swayed by the difference of variables in the argument as they attempted to solve the problem?

References

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics, 23*(3), 247-285.

Brown, S. A. (2005). The trigonometric connection: Students' understanding of sine and cosine. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (Order No. 3233908)

Dubinsky, E. (2000). Using a theory of learning in college mathematics courses. *MSOR Connections, 1*(2), 10-15.

Dubinsky, E., & McDonald, M. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton (Ed.), *The Teaching and Learning of Mathematics at University Level: An ICMI Study* (pp. 275-282). Dordrecht: Kluwer Academic Publishers.

Gray, S. S., Loud, B. J., & Sokolowski, C. P. (2009). Calculus students' use and interpretation of variables: Algebraic vs. arithmetic thinking. *Canadian Journal of Science, Mathematics and Technology Education, 9*(2), 59-72.

Küchemann, D. (1978). Children's understanding of numerical variables. *Mathematics in School, 7*(4), 23-26.

Moore, K. C. (2014). Quantitative reasoning and the sine function: The case of Zac. *Journal for Research in Mathematics Education, 45*(1), 102-138.

Phillipp, R. A. (1992). The many uses of algebraic variables. *The Mathematics Teacher, 85*(7), 557-561.

Weber, K. (2005). Students' understanding of trigonometric functions. *Mathematics Education Research Journal, 17*(3), 91-112.

27

THE DUALITY PRINCIPLE AND LEARNING TRAJECTORIES IN MATHEMATICS EDUCATION RESEARCH

Eric Weber
Oregon State University
Eric.Weber@oregonstate.edu

Elise Lockwood
Oregon State University
elise314@gmail.com

The purpose of this paper is to argue that attention to students' ways of thinking should complement a focus on students' understanding of specific mathematical content, and that attention to these issues can be leveraged to model the development of mathematical knowledge over time using learning trajectories. To illustrate the importance of ways of thinking, we draw on Harel's (2008a, 2008b) description of mathematical knowledge as comprised of ways of thinking and ways of understanding. We use data to illustrate the explanatory and descriptive power that attention to the duality of ways of understanding and ways of thinking provides, and we propose suggestions for constructing learning trajectories in mathematics education research.

Keywords: Ways of thinking, Ways of understanding, Duality, Mathematical knowledge, Learning trajectories

Introduction

In recent years, research that looks deeply at student learning has flourished, and research that focuses on the development of learning trajectories has received particular attention. Much of this research has taken substantial strides in articulating students' mathematical knowledge about particular content areas such as fractions (Simon & Tzur, 2004), partitioning and splitting (Confrey, Maloney, Nguyen, Mojica, & Myers, 2009), and length measurement (Barrett et al., 2012; Sarama, Clements, Barrett, Van Dine, & McDonel, 2011; Szilagyi, Clements, & Sarama, 2013), allowing researchers and teachers to gain much insight about students' thinking about such topics. Our reading of the work cited above, as well as our own efforts to document student thinking, suggest an important distinction between two aspects of mathematical knowledge – first, knowledge about mathematical content (i.e., knowledge of a particular proof) in particular, and second, broader aspects about that content knowledge (i.e., knowledge of what constitutes a proof) that supersede, but also interact with, knowledge about particular mathematical concepts.

We contend that both of these dimensions of mathematical knowledge should be incorporated into research involving learning, and, more specifically, learning trajectories. However, we think that incorporating both dimensions necessarily entails considering them as reflexive, and that the reflexivity between them provides a means to explain conceptual change in learning trajectories. In this paper, we propose that Harel's (2008a, 2008b, 2008c) Duality Principle is a useful theoretical lens through which to consider this relationship.

Motivation for Duality in Learning Trajectories

The distinction between content knowledge and broader aspects about that content knowledge has received attention in mathematics education. For instance, the CCSSM distinguishes between standards of mathematical content and standards of mathematical practice, (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Cuoco, Goldenberg, and Mark (1996) talk about "habits of mind" related to doing mathematics, and Harel's (2008a, 2008b, 2008c) ways of thinking and ways of understanding represent an important distinction as well.

There is some evidence of this distinction being operationalized by those focusing on learning trajectories. As an example, Ellis, Ozgur, Kulow, Dogan, and Williams (2013) describe a learning trajectory for exponential growth that focuses on students' notions of function. In this learning trajectory, Ellis et al. attend to the learning of particular mathematical content, such as the fact that rate of change of an exponential function is proportional to the amount of change in the function's argument, and they suggests sequences of tasks that help students understand such content knowledge. They complement the focus on content by proposing two fundamental views of exponential growth, which they identify as a correspondence view and a covariation view in their learning trajectory. Each of these views represents not only a particular aspect of content knowledge, but also suggests broad characteristics of how students approach exponential growth. At the same time the development of the covariation and correspondence views seem qualitatively different from the development of content knowledge such as the meaning of the base in an exponential function. Indeed, they seem to be broader aspects about content knowledge that appear repeatedly in the learning of exponential growth. Ellis et al.'s learning trajectory is an example of how students' knowledge might develop over time, where that development entails two aspects of learning: the learning of content knowledge (i.e. y increases multiplicatively by the growth factor for a unit change in x) and broader aspects about that content knowledge (i.e., a covariation or correspondence view).

The notion that of mathematical knowledge may consist of both content and broader aspects about content knowledge is not new, but consideration of how more explicit attention to mathematical this distinction might affect the development of learning trajectories has not yet been addressed. The importance of this distinction has been noted by Empson (2011), who characterized the challenges of creating a trajectory about mathematical practices.

> Most, if not all, current characterizations of learning trajectories do not address the practices that engender the development of concepts – although it's worth thinking about alternative ways to characterize curriculum standards and learning trajectories that draw teachers' attention to specific aspects of students' mathematical practices as well as the content that might be the aim of that practice (Empson, 2011, p. 573)

These examples illustrate two potential directions in which existing research on learning trajectories might be expanded. First, mathematical learning represented in trajectories often includes content and broader aspects about that content, though often the focus on broader aspects is implicit and could be made more explicit. Second, trajectories consist of conceptual "levels" but they tend not to address the learning that takes place between those levels or what mechanism(s) drives that learning. We believe we can shed light on both of these issues with a) more explicit attention to how students come to develop knowledge about the broader aspects about mathematical ideas, and b) focus on knowledge about the broader aspects of content knowledge in a way that considers its reflexive relationship with content knowledge.

To accomplish these aims, we draw on Harel's (2008a, 2008b, 2008c) description of mathematical knowledge as represented by the dual constructs of *ways of thinking* and *ways of understanding* (defined momentarily), each of which influences the other, to characterize students and experts' thinking about mathematics. By using Harel's framework, which emphasizes the reflexive relationship and interaction between these two aspects of mathematical knowledge, we argue that this perspective carries particular benefit for researchers designing and developing learning trajectories. Our reason for focusing on Harel's characterization of this distinction in mathematical knowledge we made earlier is because we feel there is power in the duality he describes. The duality captured in the relationship between ways of thinking and ways of understanding helps to uncover new insights for researchers about students' conceptual change, and these insights can lead to learning trajectories explain the nuances of how students move between various conceptual levels.

In Section 3, we characterize Harel's (2008a, 2008b) Duality-Necessity-Repeated Reasoning (DNR) framework and his definition of mathematical knowledge. His work provides the motivation for the focus on ways of thinking, and we frame our proposed contribution to learning trajectories in terms of his ideas. Then, in Section 4, we describe learning trajectories and their current roles in mathematics education research, with the aim of identifying how attention to the duality principle may provide a more complete and accurate picture of the development of mathematical knowledge. Next, in Section 5, we present a data example that highlight this duality and serves to clarify and reinforce that ways of thinking and ways of understanding develop in a reflexive manner. In Section 6, we articulate how the DNR framework provides a way to represent the cause and process of conceptual change in a learning trajectory. Finally, in Section 7, we describe ways in which researchers can make use of our recommendation to attend to the duality principle as they construct and revise learning trajectories.

Part 1: DNR Framework and Mathematical Knowledge

Harel (2008a, 2008b, 2008c) proposed the DNR based instruction framework as a way to think about the learning and teaching of mathematics. He identified the constituent parts of the framework as the *duality principle* (D), the *necessity principle* (N) and the *repeated reasoning principle* (R) that together comprise effective and meaningful mathematics instruction.

In his DNR framework, Harel (2008a) articulated the notion of a *mental act*, which includes activities like interpreting, conjecturing, explaining, searching, and problem solving (p. 3). Harel defined mathematical knowledge as consisting of *both* cognitive products of a mental act *and* the characteristics of those mental acts. He proposed that a *way of understanding* is "a particular cognitive product of a mental act carried out by an individual" (p. 4). He described a *way of thinking* as 'a cognitive characteristic of a person's ways of understanding associated with a particular mental act" (Harel, 2008a). Harel's analogy was that ways of understanding correspond to subject matter knowledge and ways of thinking correspond to conceptual tools. He articulated that mathematics consists of all the ways of understanding and all the ways of thinking that have evolved throughout history. He proposed the *duality principle* as a means of characterizing the interdependency between ways of thinking and ways of understanding.

As an example of his definition of mathematics and how duality rests upon that definition, consider the mental acts of proving and problem solving. A particular proof of a given statement is a way of understanding, whereas a proof scheme is a way of thinking. In problem solving, a solution to a particular problem represents a way of understanding, but a general problem solving strategy, applicable across a variety of problems, is a way of thinking. As mentioned above, ways of thinking are the overarching characteristics of a mental act, whereas a ways of understanding are products of that act.

A foundation of Harel's model of mathematical knowledge, illustrated in the duality principle, is that thinking and understanding are reflexive. That is, "Students develop ways of thinking through the production of ways of understanding, and, conversely, the ways of understanding they produce are impacted by the ways of thinking they possess" (Harel, 2008a). We think of this as a feedback loop, and this feedback between ways of thinking and ways of understanding is the core of the *duality principle*. Given this perspective about the learning of mathematics, a representation of learning over time (such as a learning trajectory) should articulate both aspects of that development and their reflexivity. Some existing learning trajectories focus on constructs akin to ways of thinking and understanding and make the distinction explicit (Clements & Sarama, 2009). However, no learning trajectory work of which we are aware has explicitly considered the reflexivity of ways of thinking and understanding and how the feedback loop between the spurs the development of students' knowledge.

An Introduction to Learning Trajectories in Mathematics Education

When we describe learning trajectories, we mean representations (either predictive or descriptive) of the development of students' mathematical knowledge over time. Simon and Tzur (2004) first identified a hypothetical learning trajectory (HLT) as a model of how students' learning might occur over a period of time, with particular attention paid to students' mathematical activity and the role of tasks in engendering that activity. We want to note that the construct of a learning trajectory is broader than the HLT, and in this paper we specify whether we mean LT or HLT throughout the paper. They proposed four principles for the hypothetical learning trajectory construct (Simon & Tzur, 2004, p. 93).

1) Generation of an HLT is based on understanding of the current knowledge of the students involved.
2) An HLT is a vehicle for planning learning of particular mathematical concepts.
3) Mathematical tasks provide tools for promoting learning of particular mathematical concepts and are, therefore, a key part of the instructional process.
4) Because of the hypothetical and inherently uncertain nature of this process, the teacher[1] is regularly involved in modifying every aspect of the HLT.

These principles highlight key components of learning trajectories in existing literature. First, they build on what a student understands, which requires determination of existing mathematical knowledge. Second, the HLT models how a student might develop a mathematical understanding by engaging with tasks that promote specific mathematical activity. In describing the impact of Simon's work, Duschl (2011) identified the HLT construct as the beginning of a movement that resulted in 'the recommendation that science/math learning be connected through longer sequences of instruction that function vertically across grades/years and horizontally within a given school year' (Duschl, 2011, p. 124)[2].

4.1 Common Elements Among Learning Trajectories in Mathematics Education

Researchers typically frame a learning trajectory as hypothesis generating (via an HLT as previously mentioned, (Simon & Tzur, 2004)) or hypothesis testing (via an emergent learning trajectory, defined below), though these two processes can be thought about as an iterative cycle. Examples include concepts such as graphs of multivariable functions (Weber, 2012), fractions (Saenz-Ludlow, 1994; Simon & Tzur, 2004), trigonometry (Moore, 2010), and geometric figures (Clements, Wilson, & Sarama, 2004)[3]. Each of these learning trajectories focuses on one or more specific mathematical understanding(s), proposes the mathematical knowledge students need to have a coherent view of that idea, and describes a sequence of activities and instruction to engage students in learning the idea in the way the researcher proposed. First, the researcher identifies the *hypothetical* learning trajectory (by Simon, 1994), consisting of students'

---

[1] We consider this principle as equally applicable to teacher-researchers.

[2] We note that we are discussing LTs here, and our discussion should not be confused with LP's, which we view as a theoretically distinct construct.

[3] This is not intended to be an exhaustive list, but is representative of the various ways in which LTs are being used to represent learning about a variety of mathematical topics.

mathematical knowledge and tasks designed to support that knowledge during the design phase of the experiment. Second, the researcher generates what we call the *emergent* learning trajectory (ELT) that reflects the development of the student's mathematical knowledge after actual instruction and data interpretation (Figure 1). (This has been called the *actual* learning trajectory ((Middleton, Flores, Carlson, Baek, & Atkinson, 2003). We choose to characterize it as emergent because we claim that no learning trajectory is "actual" as it represents a model, and therefore a representation, of student's thinking; also see Figure 2).

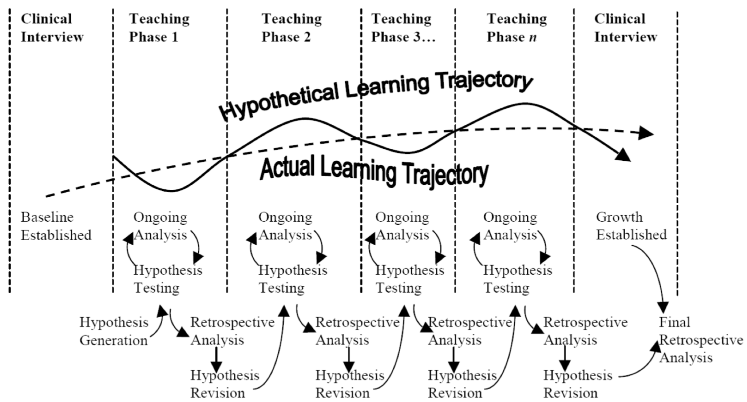


Figure 1. Middleton et al.'s (2003) characterization of constructing learning trajectories.

The *hypothetical* learning trajectory (Simon & Tzur, 2004) typically consists of a conceptual analysis of a mathematical idea and a series of instructional tasks intended to engender the ideas defined in the conceptual analysis. Thompson (2008) defined a conceptual analysis as, 'ways of thinking[4] that, if students had them, might be propitious for building more powerful ways to deal mathematically with their environments than they would build otherwise' (Thompson, 2008, p. 58). Thompson described conceptual analysis as useful in two ways. First, one can generate models of thinking that aid in explaining observed behaviors and actions of students (part of an *emergent* learning trajectory). Second, one can construct ways of understanding that, were a student to have them, might be useful for his or her development of a scheme of meanings that would constitute a coherent conception of a mathematical idea. Additionally, the development of a student's mathematical knowledge typically results from analysis of data from one or more teaching experiments. Steffe and Thompson (2000) proposed that the purpose of a teaching experiment is to experience students' mathematical learning and reasoning as a first step to constructing models of students' ways of thinking[5].

The Affordances of Explicit Attention to Duality in Learning Trajectories

Learning trajectories, as Simon characterizes them, have been used by a variety of researchers. Such uses of learning trajectories have ranged from modeling student learning about mathematical ideas (Castillo-Garsow, 2010; Moore, 2010) to serving as a tool for teachers to think about their instruction (Sztajn, Confrey, Wilson, & Edgington, 2012). We have found that

[4] Note that Thompson's use of *ways of thinking* does not reflect the distinction Harel makes with ways of understanding. Instead, we take Thompson to mean specific understandings of mathematics and refer to it as such here.

[5] The context of their use of way of thinking suggests attention to mathematical content knowledge, which for Harel is a way of understanding.

most learning trajectories tend to focus on representing the development of knowledge about specific mathematical content, and we would argue that such learning trajectories do not seem to give *explicit* attention to ways of thinking described by Harel, even when the ways of thinking might be implicit in the trajectory. While these ideas may be implicit, not identifying them constrains discussion of how ways of thinking might affect the ways of understanding students appear to use, and vice versa. For example, we recall our earlier discussion of Ellis et al.'s learning trajectory for exponential growth. Suppose that the learning trajectory only consisted of specific understandings (i.e. the function gets bigger over time) without explicit attention to the covariation or correspondence view. An LT constructed solely of ways of understanding may not recognize the power of broader, orienting approaches (such as a correspondence view) that drive the observed way of understanding. This type of trajectory also does not allow for consideration that the application of a way of thinking to a particular situation (which we see as a way of understanding) affects the structure of that way of thinking, and that conceptual changes may arise from this interaction. In Ellis et al.'s LT, for example, this could mean that as a student with a correspondence view applies that way of thinking to a particular task. Perhaps the student encounters a task that requires conceiving of continuous variation between quantities. The student's attempt to make sense of this situation with a correspondence about a necessary limitation of their correspondence view and may adjust that way of thinking accordingly. This experience, which represents a way of understanding, might in turn affect (by strengthening, putting into question, or altering) the initial steps in the development of a covariation view (or way of thinking). In this way, attention to the reflexivity of ways of thinking and ways of understanding complements the careful work about student thinking already being done.

We now can reformulate our argument, given the introduction of specific terminology. In the remainder of this paper, we seek to leverage Harel's work (Harel, 2008a, 2008b, 2008c; Harel & Koichu, 2010), particularly his characterization of the duality of *ways of thinking* and *ways of understanding*, as we propose recommendations for constructing and evaluating learning trajectories. We now present two examples of data that highlight the value of duality in considering students' mathematical knowledge. We argue that certain ways of thinking displayed in these examples complement, and at times inform, the ways of understanding that students have about particular situations, and similarly that the ways of understanding can feed back into ways of thinking. Ultimately we claim that considering both ways of understanding and ways of thinking enhances our understanding of students' mathematical knowledge.

Attending to Duality to Explain Conceptual Change: An Example from Data

In this section, we use an example to demonstrate how Harel's duality principle provides a means to explain students' conceptual change. We use previous work from a study on combinatorics to exemplify the importance of considering ways of thinking. The point of this example is to advocate that the reflexive relationship between ways of thinking and ways of understanding (and the subsequent feedback loop between the two) can help explain students' conceptual development, which can be represented using a learning trajectory. By *explain*, we mean illuminate plausible mechanisms by which change in the students' knowledge might occur. In the following discussion, it is important to keep in mind the distinction between mental acts and ways of thinking. As Harel (2008c) points out (p. 3), mental acts include activities such as proving, explaining, generalizing, problem solving, and justifying. Ways of understanding are products of such mental acts (e.g. a particular solution), while ways of thinking are cognitive characteristics of them (e.g. a particular problem solving approach like solving smaller, similar problems) (p. 4).

Data Example: Solving Smaller, Similar Problems in Combinatorics

Our example draws on Lockwood's (2013, in press) work with post-secondary students who solved five advanced counting problems in videotaped, semi-structured interviews. In this study, students solved the problems on their own and were later presented with alternative hypothetical answers (sometimes correct, sometimes incorrect), putting the students in situations in which they were comparing two possible answers. The aim was to see what students would think about and do in such situations, especially targeting how students would draw upon sets of outcomes (Lockwood, 2013). In addition to highlighting duality in another domain (combinatorics), this example shows how a student's way of thinking can span multiple ways of understanding particular problems. While Lockwood has previously framed this work as having specific content implications in a combinatorial setting (Lockwood, in press) and in terms of sets of outcomes (Lockwood, 2013), the purpose of this example is to emphasize a particular problem-solving approach (solving smaller, similar problems) as a way of thinking, and to show how that way of thinking affected both the student's ways of understanding particular problems and the researcher's interpretation of the students' combinatorial thinking. This data demonstrates a distinction Harel (2008a) himself made between instances of solving a problem (representing ways of understanding) and broader problem-solving approaches (representing ways of thinking). Indeed, he identifies "looking for a simpler problem" (p. 6) as an example of a way of thinking about the mental act of problem solving.

For Lockwood (in press) solving smaller, similar problems refers to the problem-solving approach of attempting one or more simpler versions of a problem as a means of gaining insight into a solution technique that may apply to the original problem. In any given counting problem, there are a number of conditions that determine what the problem is asking. Some of these might be numerical in nature (e.g., the specific number of letters in a password), which we call parameters, but others might refer to non-numerical conditions (e.g., the fact that repetition of letters is allowed in a password), which we call constraints. A smaller, similar problem reduces numerical parameters but maintains the constraints of the additional problem. This way of thinking is found elsewhere in the literature, typically presented as a valuable problem solving heuristic (Polya, 1957; Schoenfeld, 1979; Silver, 1981). The way of thinking of solving smaller, similar problems is demonstrated across two examples for a particular student. Anderson worked on the Passwords problem and the Groups of Students problem, which are presented in respective subsections below.

The Passwords problem. The Passwords problem states, *A password consists of eight upper case-letters. How many such 8-letter passwords contain at least three Es?* In this episode, Anderson was in the process of comparing two expressions, although he did not know which answer was correct.[6] The correct expression he was evaluating, referred to as Expression PC (for Passwords Correct) is

$$\sum_{k=3}^{8}\binom{8}{k}25^{8-k}=\binom{8}{3}\cdot 25^5+\binom{8}{4}\cdot 25^4+\binom{8}{5}\cdot 25^3+\binom{8}{6}\cdot 25^2+\binom{8}{7}\cdot 25^1+\binom{8}{8}\cdot 25^0.$$

The incorrect expression is $\binom{8}{3}\cdot 26^5$, subsequently referred to as Expression PI (for Passwords Incorrect). Anderson numerically computed both Expression PC and PI and found there to be a

[6] This activity of evaluating the expression of an alternative answer was part of the design of the interviews. The aim was to put students in situations where they had to evaluate incorrect but seemingly reasonable answers.

large numerical difference, indicating to him that one of the expressions was incorrect. As Anderson thought more about why the discrepancy occurred, he decided to truncate the situation to examining the number of 4-letter passwords that contain at least three Es, rather than 8-letter passwords that contain at least three Es (see Figure 2). In reasoning through this 4-letter smaller problem, he noted that there were 25 options for any placement of three Es, and he explained again that there were $\binom{4}{3}$ ways to place three Es in the four slots, with 25 choices for the remaining letter. He then considered the case in which all four letters were Es, and he stated that there was just one way to do that, giving him an initial answer of 101. Then, Anderson tried the 4-letter problem with Expression PI. When he did this, he arrived at $\binom{4}{3}\cdot 26$, which gave an initial answer of 104. Expression PC had given him an answer of 101, and thus he noted, "the difference is already there." We think it is significant that Anderson realized that the discrepancy between the expressions existed even in the smaller problem, and as a result he focused very closely on the small case. This numerical difference of 101 and 104 was small enough for him to be able to consider in detail, and he proceeded to examine the difference here more closely.

As he examined the smaller case, he said, "I have E, E, E, A through Z. Which is equal to 26" (while writing down E E E A-Z). Then, he said, "Then I have another E, E, A through Z, E, which is another 26" (while writing E E A-Z E). He then noted, "And since I do this four times, I have 4 times 26, which is 104, okay, which would suggest that the second one (Expression PI) is correct." Then, nearby, he wrote out E E E A-Z-E, and he said, "I have 3 Es, then I set them to any 25 letters, so let's see, A through Z minus E. And so I have 100 different ways to do that." Then he wrote E E E E and noted "But then I also have 4 Es, and there's only one way to do that." His reflection on this discovery is seen below.

Excerpt 4. Anderson recognizes the reason behind the overcounting.

A: Oh, there we go, that's where the difference is…there's 26 different ways to arrange it so that the first 3 letters are Es, and then the last one can be any of the 26 letters. And then there's another way to arrange it so that the first 2 and the last letter are Es, and the 3rd letter is any letter between A and Z, except if the third letter is an E, it's exactly, it's the exact same case as if the E was the last letter in the first case, which means it's counting multiple passwords twice.

Anderson thus identified a particular password (the all Es password) that was counted too many times by the incorrect Expression PI. After this discussion about the smaller case, Anderson was able to use his work to make sense of which expression was correct in the original problem. His use of the smaller, similar problem was a vital part of him successfully evaluating the alternative solution and determining an accurate answer. In this problem, we contend that Anderson used a way of thinking – the problem-solving approach of solving a smaller, similar problem – in order to think more about the discrepancy between the two answers.

After this episode, Anderson was asked to reflect on his use of the smaller case. We believe that his response highlights the duality principle – that his way of thinking had a direct bearing on how he conceptualized the problem and influenced how he thought about the instance of overcounting that was occurring. In his reflection, it seems that he had anticipated how working with the smaller problem would facilitate his manipulation of the passwords.

Excerpt 5. Anderson's reflection on the problem suggests a way of thinking.

E: Uh, do you feel like specific numbers there, like, knowing the 104 versus the 101 helped as opposed to having those huge numbers?

A: Yeah, um, I realized if I'm going to do this, like if I'm going to compare how I get to these numbers for an 8-letter thing it's going to be way too big. And it's going to take me a lot of time, so I was like, well, let's just start off with 4, since it gives me room, um, since 3 of them have to be E's, so I can just manipulate the last one however I wanted… And then yeah I calculated them to see if they were still different at that small of a level, here, if I started to shrink the problem. And I was like, oh, the problem – the two methods still come up with different answers, so something must be off on some fundamental level somewhere.

E: Okay.

A: So I realized, well, since my brain's not all that math oriented, I guess I'll just like write it out and see where I go, so let's come up with a few examples, so I was like EEE, and I was like, well, my brain's too lazy to come up with a specific example, so I guess I'll just write down the range, and then I should be okay. And I guess it's that step that my brain kept skipping due to laziness, (chuckles) that made me overlook this, that one problem.

The work with the smaller problem clearly helped Anderson make sense of the situation and solve the problem correctly. While Anderson never said so explicitly (and he was not asked), because he was successful on this problem, we would speculate that the way of understanding the occurrence of overcounting fed back positively into his way of thinking. That is, his work with the smaller case on the Passwords problem ended positively in a correct (and robust) understanding of why overcounting occurred. This experience likely solidified this way of thinking, encouraging him to continue to utilize smaller, similar problems in subsequent work (and indeed, we will see that he used this strategy again on the following Groups of Students problem). When analyzing his combinatorial thinking, then, and in mapping out a learning trajectory for the concepts that facilitate the solving of combinatorial problems, Anderson's way of thinking with smaller, similar problems likely affected his understanding of these concepts. When solving counting problems, part of his conceptualization is that he can adjust the size of the numbers to make a more tractable problem. Since he had success on the Passwords problem in being able to identify key mathematical ideas in the smaller problems (such as how exactly the mechanism for the overcounting was occurring), that shift to the smaller problem, we argue, became an integral part of his conceptualization of solving counting problems.

The Groups of Students problem. Later in the interview, Anderson solved the Groups of Students problem, which states, *In how many ways can you split a class of 20 into four groups of five?* A correct way to answer this problem is to choose 5 students from 20 to form a group, then 5 from 15, then 5 from 10, and finally 5 from the remaining 5, and there are $\binom{20}{5}\binom{15}{5}\binom{10}{5}\binom{5}{5}$ ways to do this. This process has an implicit order to the stages and counts the total number of distinguishable groups, and we thus divide this expression by 4! for the final answer. We briefly mention Anderson's work on this subsequent problem to show how the same way of thinking emerged again in his work on a new counting problem. This supports our above assertion that his positive experience with the way of thinking on the Passwords problem might have encouraged him to employ it subsequently. It also helps to establish his problem solving approach as a way of thinking, not just a way of understanding that he applied to a particular situation in the Passwords problem. However, we also want to use this example to show how his experience on the Groups of Students problem raised new issues in his way of thinking, thus affecting his conceptualization of how he might apply this way of thinking to novel counting problems.

In this problem, Anderson reduced two parameters (the number of groups and the total number of students) to make the problem more tractable. He divided a class of four into two groups, and through systematic listing found that there were three ways to do this. He then attempted to determine how a class of six could be split into two groups, and again through systematic listing found that there were 10 such possibilities. Anderson continued in this way, he made an initial guess at what the general formula might be: "the number of students choose the size of the groups, divided by the number of groups." We note that this formula is incorrect, but given his work it is a reasonable first attempt. Recognizing that he wanted to test out this guess at a formula, Anderson proceeded to solve another smaller problem, this time splitting six students into three groups of two. He wrote out solutions and similarly developed a pattern, continuing to reason about the problem. We ultimately ran out of time for Anderson to come up with a correct solution on his own, but his work with the smaller problems proved fruitful for him, and he was able to make sense of the correct answer, $\frac{\binom{20}{5}\binom{15}{5}\binom{10}{5}\binom{5}{5}}{4!}$.

In this example, we want to emphasize that multiple applications of the way of thinking actually might have served to refine Anderson's way of thinking of using smaller, similar problems. Anderson's way of using smaller cases in this problem differed from (and was not quite as straightforward as) his work on the Passwords problem. In the Passwords problem, insight about overcounting came to him right away, and one instance of a smaller case (reducing the problem to 4 instead of 8 letters) was enough for him to make sense of the correct solution. In the Groups of Students problem, he used a series of smaller problems to build up a pattern from which he made a guess at what the correct formula might be, and this initial formula was not correct. The nature of the interviews did not allow for explicit reflection about this way of thinking, but we can make a couple of inferences about how his ways of understanding in the two problems fed back into his way of thinking. We contend that the use of multiple smaller problems and the emergence of patterns supplemented and expanded his previous way of thinking about the mental act of problem solving. Additionally, we argue that as a result of his work on the Groups of Students problem, Anderson might have learned that he had to be careful in his choice about how to reduce the problem – simply reducing any parameters might not be helpful, and it is important for him to be strategic about how he reduces a problem. This is seen in his first unsuccessful attempt at breaking a group of eight into four groups of two, and this is an insight that might not have arisen had he only solved the Passwords problem. We would thus argue that Anderson's way of thinking is more robust because of the ways of understanding with which he engaged on these two problems.

Lockwood (in press, 2013) has made the case for combinatorial implications of this work and for the value of using smaller cases in counting. However, we emphasize that beyond the content, these two episodes together reveal an important aspect of Anderson's learning – his use of a particular problem-solving approach – that can be described in terms of a way of thinking. Even more, Anderson's work on both of these problems provides further evidence of the usefulness of the duality principle, suggesting that his ways of thinking and ways of understandings reflexively interacted as he solved counting problems. We argue that if researchers seek to articulate aspects of students' learning via learning trajectories, there could be value in targeting both ways of understanding and ways of thinking such as those that Anderson displayed. By emphasizing this duality, we suggest that the interaction between ways of understanding and ways of thinking actually shed light on Anderson's combinatorial conceptions, and these provide explanatory aspects of his work that would otherwise not arise. By observing his ways of thinking, ways of understanding, and their interaction across multiple problems, we have a more complete picture of how he thinks about counting problems.

Duality, Necessity and Repeated Reasoning as a Basis for Explaining Conceptual Change

In Part 2, we argued that many existing learning trajectories often identify discrete levels of conceptual change, but we suggested that these could be complemented with explicit attention to how learning occurs between those levels. In Part 3, we used an example to illustrate how attention to duality provides a means to think about the interaction between ways of thinking and ways of understanding in observed conceptual change. In this section, we extend these discussions to suggest that duality provides a means to elaborate what Simon et al. (2010) called "subtle shifts in thinking." (p.84) Simon proposed the need for understanding these subtle shifts to complement the work of Steffe, Cobb and Thompson (among others) that focused on hierarchies of student understandings. He argued that the focus on such hierarchies of student understandings do not sufficiently entail an explicit elaboration of the learning process; these hierarchies consist of a sequence of schemes that Simon believed did not describe what learning occurred "between" those hierarchies. We propose that duality, when combined with other elements of Harel's DNR framework, provides a means to study, and represent in an LT, the subtle shifts that Simon identified.

To illustrate this claim, we first return to Harel's (2008) DNR framework. He argued that if feedback between ways of thinking and ways of understanding was to produce lasting changes in mathematical knowledge, that lasting knowledge relied on *necessity* and *repeated reasoning* (Harel, 2008, p. 20). The *necessity principle* states that, "for students to learn what we intend to teach them, they must have a need for it, where by need is meant intellectual need, not social or economic need" (Harel, 2008b, p. 19). The crux of the necessity principle is that a student is placed in a problematic situation that s/he genuinely sees to solve (intellectual need) and the desired concept is necessary to develop in order to solve the problem. Similarly, retaining new knowledge over time relies on the students' internalization, retaining, and organization of that knowledge. The core of this learning over time is the *repeated reasoning principle*, which states, "Students must practice reasoning in order to internalize desirable ways of understanding and ways of thinking" (Harel, 2008b, p. 900). If students are to develop ways of understanding and ways of thinking that persist over time, the students must engage in problems that necessitate their development, and they must do so frequently. The repeated reasoning principle also distinguishes reasoning from practice. In other words, it is not just drill or practice of procedures. Students are engaged in repeated episodes of *reasoning* through problems. We interpret the *necessity* and *repeated reasoning* principles as mechanisms that drive the feedback loop between ways of thinking and ways of understanding.

Stepping back, what we are proposing is that the necessity and repeated reasoning principles engender the feedback loop between ways of thinking and ways of understanding that we think is essential to explaining the kinds of subtle conceptual shifts that Simon described. As an example, imagine that a researcher was interested in explaining the conceptual change that occurred between a student thinking about a graph as static and a graph as a representation of covariation. First, the student must have a *need* to develop another way of looking at a graph. However, we do not mean that the static view of a graph "disappears" when it is insufficient to solve just one problem. Instead, a notion of a graph as a representation of covariation could arise as a result of conceiving of covariation as necessary to solving a certain problem, while a static view of a graph solves others. We think that from Harel's point of view, the shifts between schemes occurs when the student *repeatedly* conceives of a covariation view of function as useful to solving problems with graphs while he or she also conceives of the static view as insufficient to do so. Implicit in our descriptions of intellectual need and repeated reasoning is that the student's initial way of thinking (static view) resulted in ways of understanding that were not sufficient to solve a particular problem, and that spurred a need for subtle shifts in the way of thinking about a graph. When these subtle shifts occur repeatedly they might produce an observable conceptual change such as shift from a static view of a graph to a covariation view of a graph. Even if learning trajectories only represent the macro-shifts, it is of great importance that the researcher considers how to engender the micro-shifts that produce them. The duality principle may help explain both the micro-shifts in conceptual knowledge (which may be so subtle they are not necessarily observable) and the macro-shifts in conceptual knowledge (which are clearly observable)

Incorporation of Duality in Designing Learning Trajectories

The purpose of this section is to provide specific recommendations for ways in which attention to the duality principle could shape how researchers think about and use learning trajectories. We suggest that researchers might recognize the potential that explicit attention to the duality of ways of thinking and ways of understanding might shape the understandings we might expect students to develop. To frame our recommendations, and to identify in what ways we see the focus on duality contributing to the current notion of a learning trajectory, we again consider Simon & Tzur's (2004) elements of a hypothetical learning trajectory. Under each element, we consider what a focus on duality contributes, and how researchers might practically focus on duality in the construction and revision of learning trajectories.

Table 1.

Implications of Incorporating Duality into Learning Trajectories

Principle	Considerations Duality Introduces	Recommendations for Researchers
Generation of an HLT is based on understanding of the current knowledge of the students involved	Understanding the current knowledge of students involved might entail a model of the students' ways of thinking and their ways of understanding that is grounded in the literature base and develops from interactions with the students.	Consider that difficulties or insights students appear to have could be related to their ways of thinking as well as their understanding of particular content. Ask questions across a variety of problems and domains to understand if their difficulties are content specific or involve ways of thinking.
An HLT is a vehicle for planning learning of particular mathematical concepts	An HLT is a vehicle for anticipating the development of mathematical knowledge, comprised of both ways of thinking and ways of understanding. The development of this knowledge can be represented as a feedback loop between ways of thinking and ways of understanding.	Explicitly articulate mental acts (such a problem solving, justifying, proving, explaining, etc.) and anticipate potential products of (ways of understanding) and characterizations of (ways of thinking) mental acts that might arise for students.

Mathematical tasks provide tools for promoting learning of particular mathematical concepts and are, therefore, a key part of the instructional process	Mathematical tasks provide a means to engender learning of mathematical concepts, and in doing so, provide a means to affect the development of ways of thinking. In turn, mathematical tasks that focus on engendering ways of thinking spur the development of particular subject matter knowledge.	Create tasks that help gain insight into students' ways of thinking by creating opportunities in tasks for students to reflect on their approaches and solutions *across* problems and situations. Be aware that certain mathematical domains may more or less effectively facilitate particular ways of thinking (e.g. solving smaller simpler problems)
Because of the hypothetical and inherently uncertain nature of this process, the teacher-researcher is regularly involved in modifying every aspect of the HLT	The teacher-researcher should consider modifications to the learning trajectory with both ways of thinking and ways of understanding in mind, developing and refining activities intended to engender those ways of thinking and ways of understanding.	Document the development of students' content knowledge in conjunction with their ways of thinking, explicitly attending to how these two aspects of students' mathematical knowledge interact.

We think there are three important considerations to be made going forward. First, our recommendations should not be considered as a call to develop a new "type" of learning trajectories. Instead, we have provided suggestions for what incorporating duality existing models of learning trajectories. The research on learning trajectories in mathematics education is robust and growing, but we hope our recommendations can supplement the work that is already being done. Second, our recommendations should not be considered an exhaustive list. We include these points as a means to promote discussion about a handful of important issues to consider about the inclusion of ways of thinking in learning trajectories. We hope that others may push back and refine our ideas about how to incorporate ways of thinking in learning trajectories. Third, content plays a significant role in how ways of thinking develop, and certain domains may be more appropriate than others for fostering specific ways of thinking. For instance, combinatorics is a particularly powerful context for thinking of solving smaller, similar problems, as the need for such work with smaller cases can easily be motivated in a setting that typically deals with very large and unwieldy numbers. While this way of thinking may be effectively developed in a domain like combinatorics, it could be further refined and developed in other mathematical areas, each of which might elicit different aspects of the way of thinking.

In conclusion, in this paper we have focused on how emphasizing duality might supplement the development and implementation of learning trajectories. In subsequent work, we plan to provide more empirical evidence of how duality might arise in learning trajectories. Additionally, while our argument on this paper centers on learning trajectories for particular mathematical topics (such as ratio, counting problems, or fractions), we have also wondered about the development of learning trajectories for mathematical practices (such as generalization, problem solving, or proof). The existence of this other type of learning trajectory is a theoretical question that we would like to explore further.

References

Barrett, J. E., Sarama, J., Clements, D. H., Cullen, C., McCool, J., Witkowski-Rumsey, C., & Klanderman, D. (2012). Evaluating and improving a learning trajectory for linear measurement in elementary grades 2 and 3: A longitudinal study. *Mathematical Thinking and Learning, 14*(1), 28-54.

Castillo-Garsow, C. (2010). *Teaching the Verhulst model: A teaching experiment in covariational reasoning and exponential growth*. doctoral dissertation. Arizona State University. Tempe, AZ.
Clements, D. H., & Sarama, J. A. (2009). *Learning and teaching early math: The learning trajectories approach*: Routledge.
Clements, D. H., Wilson, D. C., & Sarama, J. (2004). Young children's composition of geometric figures: A learning trajectory. *Mathematical Thinking and Learning, 6*(2), 163-184.
Confrey, J., Maloney, A., Nguyen, K., Mojica, G., & Myers, M. (2009). *Equipartitioning/splitting as a foundation of rational number reasoning using learning trajectories.* Paper presented at the 33rd Conference of the International Group for the Psychology of Mathematics Education, Thessaloniki, Greece.
Cuoco, A., Goldenberg, P., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curricula. *Journal of Mathematical Behavior, 15*, 375-402.
Duschl, R. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education, 47*(2), 123-182.
Ellis, A., Ozgur, Z., Kulow, T., Dogan, M. F., & Williams, C. (2013). *An exponential growth learning trajectory.* Paper presented at the 37th Conference of the International Group for the Psychology of Mathematics Education, Kiel, Germany.
Empson, S. B. (2011). On the idea of learning trajectories: Promises and pitfalls. *The Mathematics Enthusiast, 8*(3), 571-596.
Harel, G. (2008a). DNR perspective on mathematics curriculum and instruction, Part I: focus on proving. *ZDM Mathematics Education, 40*, 487-500.
Harel, G. (2008b). A DNR perspective on mathematics curriculum and instruction. Part II: with reference to teacher's knowledge base. *ZDM Mathematics Education, 40*, 893-907.
Harel, G. (2008c). What is mathematics? A pedagogical answer to a philosophical question. *Current issues in the philosophy of mathematics from the perspective of mathematicians*, 265-290.
Harel, G., & Koichu, B. (2010). An operational definition of learning. *The Journal of Mathematical Behavior, 29*(3), 115-124.
Lockwood, E. (2013). *Developing facility with sets of outcomes by solving smaller, simpler counting problems.* Paper presented at the 16th Annual Conference on Research in Undergraduate Mathematics Education, Denver, CO.
Lockwood, E. (in press). The strategy of solving smaller, simpler problems in the context of combinatorial enumeration. *Research in Collegiate Mathematics Education*.
Middleton, J. A., Flores, A., Carlson, M. P., Baek, J., & Atkinson, R. (2003). A longitudinal study of the development of rational number knowledge in the middle grades *Proposal to the United States National Science Foundation*.
Moore, K. C. (2010). *The role of quantitative reasoning in precalculus students learning central concepts of trigonometry*. doctoral dissertation. Arizona State University. Tempe, AZ.
National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). Common Core Standards for Mathematics.
Polya, G. (1957). *How to solve it*. Princeton, NJ: Princeton University Press.
Saenz-Ludlow, A. (1994). Michael's fraction schemes. *Journal for Research in Mathematics Education, 25*(1), 50-85.
Saldanha, L., & Thompson, P. W. (1998). Re-thinking co-variation from a quantitative perspective: Simultaneous continuous variation. In S. B. Berensah, K. R. Dawkings, M.

Blanton, W. N. Coulombe, J. Kolb, K. Norwood & L. Stiff (Eds.), *Proceedings of the Twenieth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 298-303). Columbus, OH: ERIC Clearninghouse for Science, Mathematics, and Environmental Education.

Sarama, J., Clements, D. H., Barrett, J., Van Dine, D. W., & McDonel, J. S. (2011). Evaluation of a learning trajectory for length in the early years. *ZDM, 43*(5), 667-680.

Schoenfeld, A. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education, 10*(3), 173-187.

Silver, E. A. (1981). Recall of mathematical problem information: Solving related problems. *Journal for Research in Mathematics Education, 12*(1), 54-64.

Simon, M., Saldanha, L., McClintock, E., Akar, G. K., Watanabe, T., & Zembat, I. O. (2010). A developing approach to studying students' learning through their mathematical activity. *Cognition and Instruction, 28*(1), 70-112.

Simon, M., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning, 6*(2), 91-104.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Szilagyi, J., Clements, D. H., & Sarama, J. (2013). Young childrens' understandings of length measurement: Evaluating a learning trajectory. *Journal for Research in Mathematics Education, 44*(3), 581-610.

Sztajn, P., Confrey, J., Wilson, P. H., & Edgington, C. (2012). Learning trajectory based instruction: Toward a theory of teaching *Educational Researcher, 41*, 147-156.

Thompson, P. W. (2008). Conceptual analysis of mathematical ideas: Some spadework at the foundation of mathematics education. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sepulveda (Eds.), *Proceedings of the Annual Meeting of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 45-64). Morelia, Mexico: PME.

Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education* (pp. 33-57). Laramie, WY: University of Wyoming.

Students' Struggle With the Temporal Order of Delta and Epsilon Within the Formal Definition of a Limit

Aditya P. Adiredja and Kendrice James
University of California, Berkeley

Studies about students' understanding of the formal definition of a limit, or the epsilon delta definition suggest that the temporal order of delta and epsilon is an obstacle in learning the formal definition. While such difficulty has been widely documented, patterns of students' reasoning are largely unknown. This study investigates the degree of difficulty students have with the temporal order, along with justifications that students provide to support their claim. diSessa's Knowledge in Pieces provides a suitable framework to explore the context specificity of students' knowledge as well as the potential productivity of their prior knowledge in learning.

Keywords: limit, formal definition, students' prior knowledge, fine-grained analysis

The formal definition of a limit of a function at a point, as given below, also known as the epsilon-delta definition, is an essential topic in mathematics majors' development that is introduced in calculus. We say that the limit of $f(x)$ as x approaches a is L, and write

$$\lim_{x \to a} f(x) = L$$

if and only if, for every number ε greater than zero, there exists a number δ greater than zero such that for all numbers x where $0 < |x - a| < \delta$ then $|f(x) - L| < \varepsilon$. The formal definition provides the technical details for how a limit works and introduces students to the rigor of calculus. Yet research shows that thoughtful efforts at instruction at most leave students – including intending and continuing mathematics majors – confused or with a procedural understanding about the definition (Cottrill et al., 1996; Oehrtman, 2008; Tall & Vinner, 1981).

Although studies have sufficiently documented that the formal definition is a roadblock for most students, little is known about how students actually attempt to make sense of the topic, or about the details of their difficulties. Most studies have not prioritized students' sense making processes and the productive role of their prior knowledge (Davis & Vinner, 1986; Przenioslo, 2004; Williams, 2001). This may explain why they reported minimal success with their instructional approaches (Davis & Vinner, 1986; Tall & Vinner, 1981). Thus, understanding the difficulty in the teaching and learning of the formal definition warrants a closer look – with a focus on student cognition and with attention to students' prior knowledge. It also calls for a theoretical and analytical framework that focuses on understanding the nature and role of students' intuitive knowledge in the process of learning.

A small subset of the studies have begun exploring more specifically student understanding of the formal definition (Boester, 2008; Knapp and Oehrtman, 2005; Roh, 2009; Swinyard, 2011). They suggest that students' understanding of a crucial relationship between two quantities, ε and δ within the formal definition warrants further investigation. Davis and Vinner (1986) call it the *temporal order* between ε and δ, that is the sequential ordering of ε and δ within the formal definition where ε comes first, then δ (p. 295). They found that students often neglect its important role. Swinyard (2011) found that the relationship between the two quantities is one of the most challenging aspects of the formal definition for students. Knapp and Oehrtman (2005) and Roh (2009) document this difficulty for advanced calculus students. This difficulty is also prevalent among the majority of calculus students who struggled with the formal definition in Boester (2008). How students reason about the temporal order still remains an open question.

This study is a part of a larger study investigating the role of prior knowledge in student understanding of the formal definition. It specifically explores the claim that students struggle to understand the temporal order of ε and δ within the formal definition. We aim to answer the following research questions:

1. What claims do students make about the temporal order of ε and δ?
2. If students in fact struggle with the temporal order, what is the nature of their difficulty?

Theoretical Framework

The Knowledge in Pieces (KiP) theoretical framework (Campbell, 2011; diSessa, 1993; Smith et al., 1993) argues that knowledge can be modeled as a system of diverse elements and complex connections. From this perspective uncovering the fine-grained structure of student knowledge is a major focus of investigation, and simply characterizing student knowledge as misconceptions is viewed as an uninformative endeavor (Smith et al, 1993). Knowledge elements are context-specific; the problem is often inappropriate generalization to another context (Smith et al, 1993). For example, "multiplication always makes a number bigger" is not a misconception that just needs to be removed from students' way of thinking. Although this assertion would be incorrect in the context of multiplying numbers less than 1, when applied in the context of multiplying numbers greater than 1, it would be correct. Paying attention to contexts, KiP considers this kind of intuitive knowledge a potentially productive resource in learning (Smith et al., 1993). This means that instead of focusing on efforts to replace misconceptions, KiP focuses on characterizing the knowledge elements and the mechanisms by which they are incorporated into, refined and/or elaborated to become a new conception (Smith et al., 1993). Similarly, we view students' prior knowledge as potentially productive resources for learning. We focus our investigation on students' reasoning as potentially productive resources, and we focus our attention on the context specificity of students' knowledge.

Methods

The data for this report comes from a larger interview study with 25 students (18 new students, and 7 students from the pilot study reported last year) investigating the role of prior knowledge in student understanding of the temporal order. Each of these students has received some form of instruction on the formal definition. So we anticipate some knowledge about the definition to be a part of their prior knowledge. The protocol was designed to elicit student understanding of the formal definition, but more specifically their understanding of the relationship between delta and epsilon. To explore the stability and context specificity of students' knowledge across different contexts, we asked students about the temporal order of the two variables in four different contexts: dependence (does epsilon depend on delta, vice versa or they depend on each other?), their *temporal order* (does delta or epsilon come first in the definition?), set (which one is set? Delta, epsilon, both or neither?), and lastly we asked students to order $x, f(x)$, epsilon and delta in order according to the definition. Each individual interview lasted about 2 to 3 hours. These interviews were videotaped following recommendations in Derry et al. (2010).

The first step in analysis is categorizing students' response to each question about the temporal order into the category of delta first, epsilon first or no order. The response to each question is given a score from 0 to 2 (delta first=0, no order=1, epsilon first=2). Then students' scores are added up across the different questions to give the student a total score. The total score ranges from 0 to 8 and puts students along a continuum between the claim of delta first and epsilon first. This is a more refined way of assessing students' responses about the temporal order in comparison to our pilot work when we simply relied on what the student said last about the temporal order.

The second step in the analysis is to record students' reasoning or justification for their claim about the temporal order. To do this we explore their response to the "why" question after each question about the temporal order. Common justifications emerged from the data and we were able to document the number of students who used the same justification for their claim. The goal of the analysis is not to come up with an exhaustive list of justifications for any student in calculus. Instead we aim to report the justifications that we found in the 18 students that we interviewed along with the 7 students from the pilot study.

Results

Relationship Between the ε and δ

The table below shows how each student in the study answered each question about the temporal order. The table is split into two. The top half includes students from the current study and the bottom half are students from the pilot study whose results were reported last year.

Student	Dependence	Temporal	Set	Order	Total
CL	0	0	0	0	0
SW	0	0	0	0	0
SF	0	0	0	0	0
VB	0	0	0	0	0
TF	0	0	0	0	0
JIB	0	0	1	0	1
ADH	0	0	1	0	1
JOB	0	0	1	0	1
JC	1	0	0	0	1
BP	0	0	2	0	2
BM	0	2	1	0	3
KG	0	2	1	0	3
SN	1	0	2	0	3
GA	2	0	0	2	4
RR	0	1	2	2	5
CW	1	2	1	1	5
RM	2	2	1	2	7
IL	2	2	2	2	8
DL	0	N/A	N/A	N/A	0
JJ	0	N/A	N/A	N/A	0
DR	0	0	N/A	0	0
SR	0	0	N/A	0	0
OB	0	0	N/A	0	0
AD	2	2	2	N/A	8
DC	2	N/A	N/A	N/A	8

Table 1. Students' responses to each question about the temporal order sorted from lowest to highest total and separated by current study vs. pilot study

As we reported last year, during the pilot not all of the questions were asked. AD and DC scored an 8 without answering the other three questions because they answered the questions that were asked normatively and was able to explain the formal definition accurately. To assist in parsing the table above, we charted the number of questions that students answered correctly (epsilon first, score=2).

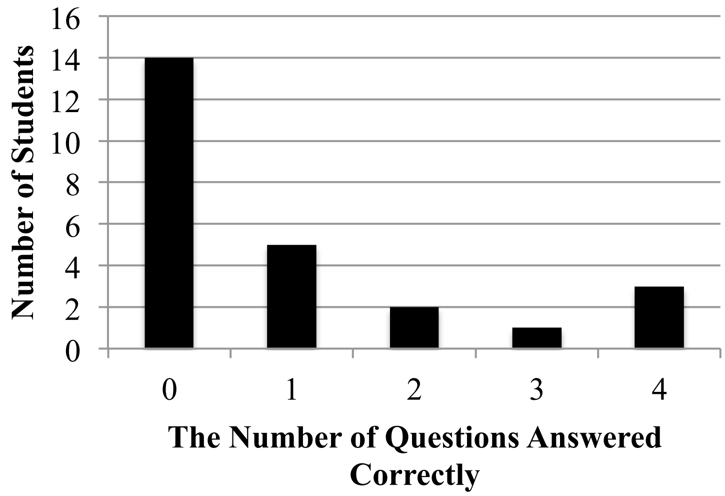


Figure 1. The distribution of students in answering the four temporal order questions correctly

Fifty six percent (56%) of students (14/25) answered none of the questions correctly, 20%(5/25) answered one question correctly, 8% answered two questions correctly, 4%(1/25) answered three questions correctly and 12% (3/25) answered four questions correctly.

Students' Reasoning About the Temporal Order

The table below shows the different justifications students provided to justify their claim about the temporal order and the number of students who used that justification. Each of the reasoning is followed by the conclusion students often make about the temporal order, along with any variations of the justification. Students from the pilot study are marked by parentheses.

Reasoning	Students	Number of students
There needs to be a delta for epsilon to exist. Often comes from students' reading of the different parts of the definition. *Variation:* the delta has to be satisfied first before epsilon could be. So delta is first.	ADH, BM, GA, KG, SN, TF, VB, (DL, JJ)	9
Delta is known because *x* and *a* are known (and L is unknown). So delta is first.	ADH, JOB, RR, VB, (SR, DR, OB)	7
Normative reading of the	BM, GA, IL, SN	4

statement. So epsilon is first.		
Spatial location of variables within the statement.	BM, GA, RR	3
Intuitive understanding of limit. Small change in x, leads to small change in $f(x)$. Small delta implies small epsilon. So delta is first.	BP, JOB	2
You're finding delta first. Delta is unknown. So delta is first.	BP, IL	2
If delta is true then epsilon is true. Often comes from students' reading of the if-then statement from the definition. So delta is first.	CL, GA, IL, JIB, KG, (JJ, SR, AD)	8
Both epsilon and delta need to be found. So no order.	CW, JC	2
Students mistaking if-then as if and only if. So no order.	JC, SW, (AD)	2
Input/output or dependence between $f(x)$ and x.	JIB, JOB, SF, SN, SW, VB, (DL, DC, JJ, AD, DR)	11
Delta is x and epsilon is y. *Variation:* y depends on epsilon and x depends on delta. Delta is grouped with x and epsilon with y.	JIB, JOB, RR, SF, SN, SW, VB, (DC, DL, JJ)	10
Partial recall from doing proofs. When done correctly leads to epsilon first, but if not, then usually delta is first.	JC, KG, RM, SF, SW, VB, (SR, DC, AD)	9
Epsilon is arbitrary, so it's not set. So delta is first.	KG, RM, SW	3
When the limit does not exist as a counter example. For a given epsilon there does not exist a delta. So epsilon is first.	(DC, AD)	2

Table 2. The list of students' justification for their claim about the temporal order.

Discussion and Implication

This study confirms the finding from our pilot study last year. We found that students struggle with the temporal order of epsilon and delta within the formal definition. Fifty six percent of the students in this study were not able to answer *one* question about the temporal order correctly. With a more refined method, we saw more variability in students' claim about the temporal order. Ten students received a total score of 0 across the four different contexts and two students scored 8, but the many students were somewhere in between. The finding that some students can score a 2 in one context but a 1 or 0 in others shows that student knowledge about the temporal order is not quite stable across the different contexts. This supports our theoretical assumption that knowledge is context specific, and also highlights the importance of assessing student knowledge in multiple contexts in research and practice.

With respect to their reasoning, *the functional dependence between x and f (x)* along with *delta is with x while epsilon is with y* remain the most common justification for the temporal order this year. We discussed the nature of that reasoning and its implication in Adiredja and James (2013). However, we also found other common reasoning patterns, like students claiming that "for every number epsilon there exists a number delta" from the definition suggests that there needs to be a delta first for epsilon to exist. We included this as one of the *knowledge resources* last year. In Adiredja and James (2013) we found that most of these knowledge resources were mathematical in nature, and hypothesized that either this indicated lack of access into the formal definition using intuitive knowledge or it was a product of using too large of a grain size to find intuitive *knowledge resources*. It turns out to be both.

The findings from this study show that for the most part students use their interpretation and experiences with formal mathematics to make sense of the temporal order. At the same time, a *microgenetic* case study as part of the larger project reveals that many of what we found in this study are reasoning patterns, and not quite yet knowledge resources. A reasoning pattern is made up of different knowledge resources, making it larger in grain size. These reasoning patterns point us in a direction to locate knowledge resources. For example, one of the authors found that delta is both a *determiner* and *a determined* in the definition by focusing on students' interpretations of the if-then statement and "for every number epsilon there exists a number delta." While delta is *determined* by epsilon, delta also *determines* the appropriate interval whose difference with the limit would later be compared to epsilon. The case study also finds that the common "gloss" *delta is x and epsilon is y* might be a byproduct of the richness and complexity of the formal definition as a learning context. That is, it might be a necessary move for student to be able to focus on the other components of the formal definition while reasoning about the temporal order.

The list that we provided here is not exhaustive, but it shows the diversity and range of student reasoning pattern. While we argue that it is too early to turn our findings into some form of instructional intervention, it is important to reiterate the point we made in Adiredja and James (2013). The goal in instruction should not be to replace some of the unproductive reasoning patterns. Instead, any instructional intervention should help students reorganize these reasoning patterns while recognizing the contexts in which they might be useful (e.g. the productivity of the functional dependence relationship in multiple contexts in mathematics). More importantly, we argue that we need to get to the level of resources to truly understand *how* students reason with the temporal order, and the ideas that they prioritize. Then we can begin to think about impactful instructions that can assist students in understanding the temporal order, and the formal definition more broadly, as a result of honoring their prior knowledge.

References

Adiredja, A.P. & James, K. (2013). Students' knowledge resources about the temporal order of delta and epsilon. In M. Zandieh (Ed.), Conference Papers Volumes in Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education, Denver, CO.

Boester, T. (2008). *A design-based case study of undergraduates' conceptions of limits*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT 3314378).

Campbell, M. E. (2011). *Modeling the co-development of strategic and conceptual knowledge in mathematical problem solving.* (Doctoral dissertation). Retrieved from http://gradworks.umi.com/3498781.pdf

Cottrill, J., Dubinsky, E., Nichols, D., Schwingendorf, K., Thomas, K., Vidakovic, D. (1996). Understanding the limit concept: Beginning with a coordinated process scheme. *Journal of Mathematical Behavior, 15,* 167-192.

Davis, R. B., & Vinner, S. (1986). The notion of limit: Some seemingly unavoidable misconception stages, *Journal of Mathematical Behavior, 5*, 281-303.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., Hall, R., Koschmann, T., Lemke, J. L., Sherin M. G. & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences, 19*(1), 3-53.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10* (2-3), 105-225.

Knapp, J., & Oehrtman, M. (2005). Temporal Order in the Formal Limit Definition. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*. Melbourne: PME.

Oehrtman, M. (2008). Layers of abstraction: Theory and design for the instruction of limit concepts. In Carlson, M. P. & Rasmussen, C. (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 65-80), *MAA Notes #73.*

Przenioslo, M. (2004). Images of the limit of function formed in the course of mathematical studies at the university. *Educational Studies in Mathematics, 55*, 103–132.

Roh, K. H. (2009). An empirical study of students' understanding of a logical structure in the definition of limit via the e-strip activity. *Educational Studies in Mathematics.* DOI: 10.1007/s10649-009-9210-4.

Schoenfeld, A. H., Smith, J., & Arcavi, A. (1993). Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. IV, pp. 55-175). Hillsdale, NJ: Erlbaum.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, ***3***(2), 115-163.

Swinyard, C. (2011). Reinventing the formal definition of limit: The case of Amy and Mike. *The Journal of Mathematical Behavior, 30,* 93-114.

Tall, D.O. & Vinner, S. (1981). Concept image and concept definition in mathematics, with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*, 151-169.

Williams, S. (2001). Predications of the limit concept: An application of repertory grids. *Journal for Research in Mathematics Education, 32*, 341-367.

THE ABILITY TO REJECT INVALID LOGICAL INFERENCES PREDICTS PROOF COMPREHENSION AND MATHEMATICS PERFORMANCE

Lara Alcock*, Toby Bailey**, Matthew Inglis*, Pamela Docherty**
* Loughborough University, UK
** University of Edinburgh, UK

In this paper we report a study designed to investigate the impact of logical reasoning ability on proof comprehension. Undergraduates beginning their study of proof-based mathematics were asked to complete a conditional reasoning task that involved deciding whether a stated conclusion follows necessarily from a statement of the form "if p *then* q*"; they were then asked to read a previously unseen proof and to complete an associated comprehension test. To investigate the broader impact of their conditional reasoning skills, we also constructed a composite measure of the participants' performance in their mathematics courses. Analyses revealed that the ability to reject invalid denial-of-the-antecedent and affirmation-of-the-consequent inferences predicted both proof comprehension and course performance, but the ability to endorse valid modus tollens inferences did not. This result adds to a growing body of research indicating that success in advanced mathematics does not require a normatively correct material interpretation of conditional statements.*

Key words: Conditional Inference, Logical Reasoning, Proof Comprehension, Undergraduate Mathematics Education

Introduction

Mathematics and logical reasoning are seen as closely related. It is widely believed that study of mathematics develops general logical reasoning skills (e.g., NCTM, 2000), and that correct logical reasoning is important for study of advanced mathematics: transition-to-proof textbooks commonly deal with the topic explicitly. But is this really the case? Recent research has revealed that mathematics and logical reasoning are related, but that this relationship is not straightforward: it is not the case that experienced mathematicians uniformly conform to normatively correct interpretations of conditional statements (e.g., Inglis & Simpson, 2006). This raises questions about what we need to teach and about which failures of reasoning should worry us. In this paper we take up this discussion, arguing that mathematics education does not lead students to normatively correct reasoning, but does nevertheless develop the logical reasoning skills that students need for advanced mathematics. We begin by reviewing arguments and evidence on expert and novice reasoning with *conditional statements* of the form "if *p* then *q*."

Reasoning with Conditional Statements

Consider a conditional statement about an imaginary letter-number pair:

"If the letter is X then the number is 1."

Researchers have investigated patterns of responses to four possible inferences from this statement plus a related assertion; these inferences are listed in Table 1.

Assertion	Inference	Inference type
The letter is X (p)	The number is 1	*Modus ponens*
The letter is not X (not-p)	The number is not 1	*Denial of the antecedent*
The number is 1 (q)	The letter is X	*Affirmation of the consequent*
The number is not 1 (not-q)	The letter is not X	*Modus tollens*

Table 1: Four possible inferences from the conditional statement plus a related assertion.

In formal logic, where conditional statements are interpreted as *material conditionals*, modus ponens (MP) and modus tollens (MT) inferences are defined to be valid, and denial-of-the-antecedent (DA) and affirmation-of-the-consequent (AC) inferences are defined to be invalid. This is the interpretation taught in textbooks and in transition-to-proof-courses. It is not, however, the interpretation commonly made by people without formal training. In everyday life, it has been argued that a *defective conditional* interpretation is more common (e.g., Quine, 1966). Under this interpretation, the conditional statement is seen as stating only that the consequent follows given that the antecedent is true, meaning that the statement is irrelevant in cases in which the antecedent is not true. The picture is further complicated by the common everyday interpretation of a conditional statement "if p then q" as the biconditional statement "p if and only if q" (e.g., Epp, 2003). These three interpretations are compared in Table 2.

p	q	if p then q (biconditional)	if p then q (defective)	if p then q (material)
T	T	T	T	T
T	F	F	F	F
F	T	F	irrelevant	T
F	F	T	irrelevant	T

Table 2: Comparison of biconditional, defective conditional and material conditional interpretations (T – true; F – false).

The corresponding responses to the four inference types are given in Table 3. Under a biconditional interpretation, p and q are seen as simply "going together," so that either both are true or both are false. This corresponds to endorsement of all four inferences. Under a defective interpretation, only modus ponens is endorsed, since the other three inferences involve assertions for which the conditional statement is seen as irrelevant. The material interpretation corresponds to the normatively correct responses as listed above.

Inference type	biconditional	defective	material
Modus ponens	endorse	endorse	endorse
Denial of the antecedent	endorse	reject	reject
Affirmation of the consequent	endorse	reject	reject
Modus tollens	endorse	reject	endorse

Table 3: Comparison of biconditional, defective conditional and material conditional interpretations.

It might be natural, then, to see the material interpretation as the most sophisticated, and to believe that mathematics educators should help students develop toward this interpretation and should be concerned if it is not attained. But is it true that a material interpretation is important for mathematical success?

Evidence on Mathematical Education and Logical Reasoning

It has been argued that a material interpretation is needed for advanced mathematics; that certain forms of indirect reasoning are not accessible without it (Durand-Guerrier, 2003). However, it has also been argued that at lower educational levels a defective interpretation is more appropriate, “since in school mathematics, students have to appreciate the consequence of an implication when the antecedent is taken to be true” (Hoyles & Küchemann, 2002, p. 196). Indeed, evidence indicates that mathematical study develops conditional reasoning skill, but develops it toward a defective rather than a material interpretation. In a sample of students in the UK, where compulsory education ends at 16, those studying mathematics in the first non-compulsory year were found to change in their responses to a conditional reasoning task more than did an equivalent population studying English literature (and not mathematics). The mathematics students became more likely to reject AC inferences and DA inferences, but also more likely to reject MT inferences (Attridge & Inglis, 2013). Might this mean that their education did a disservice to those who went on to study undergraduate mathematics? Does pre-proof mathematical education teach students a better but still inadequate interpretation of the conditional, and does this cause problems when they come to study proof?

Surprisingly, evidence from work with professional mathematicians suggests that it might not, because mathematicians do not reliably make the material interpretation either. In a study of mathematicians’ responses to the Wason Selection Task, Inglis and Simpson (2006) showed that professional mathematicians behave differently from members of a general educated population: they are not tempted by AC and DA inferences, but neither do they reliably consider a relevant MT inference. Perhaps, then, a defective interpretation is perfectly adequate for success even in proof-based mathematics. Our data supports this suggestion, as described below.

Methods

Participants in our study were 112 students in a first year, second semester undergraduate mathematics class on problem solving and proving (the equivalent of a transition-to-proof course). All had taken a linear algebra course in the previous semester (this included theorems and proofs but treated these quite lightly) and were concurrently enrolled on a course in calculus (this included some proofs and some calculations involving limits, but epsilon-delta techniques appeared only briefly). All were spending 50% of their total time over the year in these mathematics classes, and for almost all this was a compulsory component of a degree programme with “mathematics” in the title. In workshop sessions in week 8 of the 11-week problem solving and proofs course, participants were asked to complete a conditional reasoning test, and to read and answer comprehension questions on a previously unseen proof. They completed both tasks individually and in silence.

The conditional reasoning test (adapted from Evans, Clibbens & Rood, 1995) comprised 16 items of the form shown in Figure 1. There were four items for each type of inference (MP, AC, DA and MT), and instructions asked participants to decide whether the conclusion necessarily follows and to indicate their answer by placing a check mark in the appropriate circle. Participants were given ten minutes to complete the task, and the order of the items was randomised for each participant. For analysis purposes, a count out of four was constructed for each inference type, where each point indicated an instance in which the participant agreed with an inference of the relevant type.

If the letter is J then the number is not 2.
The number is 7.
Conclusion: The letter is J.
○ YES.
○ NO.

Figure 1: A conditional reasoning test item (an AC item).

The proof comprehension task involved a proof that the product of two primes is not abundant (i.e., that the product is not less than the sum of its proper factors). Participants were asked to study the proof carefully and then to answer a proof comprehension test based on the model developed by Mejía-Ramos, Fuller, Weber, Rhoads and Samkoff (2012). This test comprised ten multiple-choice items, each of which had two distractors and one correct answer (the proof and test are too long to reproduce here, but full copies will be provided at the talk if this paper is accepted). Participants were allowed 15 minutes for this task.

Finally, we obtained the participants' examination scores in their calculus, linear algebra and problem solving and proofs courses (all three courses had some coursework together with a final individual summative examination worth 85% of the course grade). The average of these scores was used as a measure for performance in core mathematics courses.

Results

Table 4 presents the descriptive statistics for all six measures, showing the minimum and maximum number of inferences of each type endorsed, as well as the associated means and standard deviations.

Measure (theoretical max)	Min	Max	Mean	Std. Dev.
MP inferences endorsed (4)	2	4	3.68	0.541
DA inferences endorsed (4)	0	4	1.06	1.085
AC inferences endorsed (4)	0	4	1.42	1.271
MT inferences endorsed (4)	0	4	2.68	1.050
Proof Comprehension (10)	3	10	7.29	0.182
Math Course (100)	22	95	61.86	16.426

Table 4: Descriptive statistics.

We note that the MP counts were close to ceiling. This is to be expected, but it renders this measure inappropriate for use in regression models, so we omit it in the following analyses. The counts for the remaining conditional reasoning measures show considerable variability – participants on average endorsed more than one of each of the invalid DA and AC inferences, and rejected more than one valid MT inference. The proof comprehension scores were generally high, and the range and average of the mathematics performance scores were typical in the national context in which the study took place.

Table 5 presents regression models with the DA, AC and MT counts as independent variables and with (a) proof comprehension score and (b) mathematics performance score as the dependent variables. Figure 3 shows the means of the proof comprehension and mathematics performance scores for participants with different DA, AC and MT counts, together with lines of best fit for cases in which the count is a significant predictor.

(a)

R^2	Predictors	β	p
.171***	DA	-.234	.029
	AC	-.234	.031
	MT	.003	.969

(b)

R^2	Predictors	β	p
.261***	DA	-.209	.041
	AC	-.342	.001
	MT	.084	.321

Table 5: Regression models predicting (a) proof comprehension score and (b) mathematics performance; ***p<.001.

.

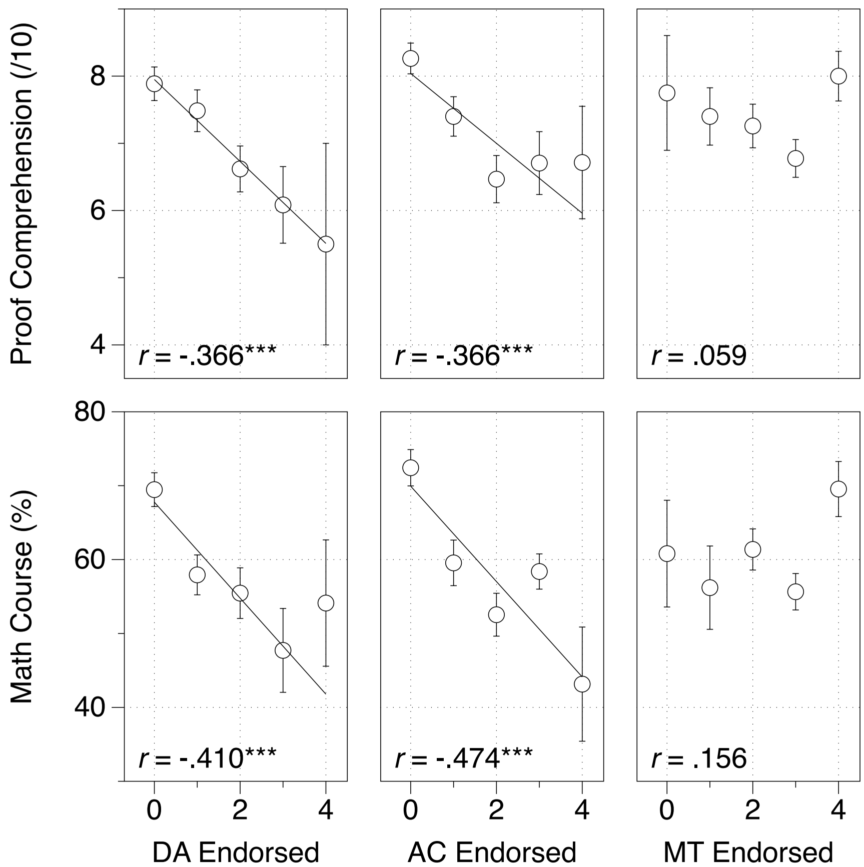


Figure 2: Proof Comprehension test and Math Course score means and correlation coefficients for participants endorsing different numbers of DA, AC and MT inferences; error bars show ±1 SE of the mean, *** $p < .001$.

In both models, the DA and AC counts were significant predictors with negative coefficients: participants who rejected more DA and more AC inferences performed better both on the proof comprehension test and in their mathematics courses. In both models, the MT score was not a significant predictor: endorsement of MT inferences, which are valid under the normatively correct material interpretation of the conditional, did not have a systematic effect on either outcome measure.

Discussion

The material interpretation of the conditional is normatively correct and is taught in standard undergraduate mathematics. However, this study adds to a growing set of results suggesting that full conformity with its entailments is not necessary for mathematical success. While the ability to reject invalid DA and AC inferences does appear to predict success in proof comprehension and in undergraduate-level courses, the ability to reliably endorse valid MT inferences does not.

One obvious limitation of this study is that it involved a comprehension test for only a single proof. Concern about this should be mitigated by the second regression model in which performance across three core mathematics modules showed a similar pattern; if a lack of ability to endorse MT inferences were a serious problem, we would expect it to appear as a significant predictor in this model. However, these results leave open the possibility that a defective interpretation of the conditional is a disadvantage under some specific circumstances. Perhaps, for instance, students with this interpretation are less able to understand contradiction or contraposition arguments. All such arguments relative to a statement of the form "if *p* then *q*" involve a step at which one establishes not-*q* and uses this to conclude not-*p*. Thus, we might expect them to be less well understood by students who do not readily endorse MT inferences. This could be investigated, although we suggest that such work should be done in parallel with further investigation of how expert mathematicians process such arguments. Recall that mathematicians do not reliably consider relevant MT inferences under all circumstances, so there might not be a straightforward link between reasoning about single abstract conditional statements and understanding this structure as it is used in proofs. Indeed, Inglis and Simpson (2009) suggest that the equivalent of an MT inference can be constructed given a defective interpretation of the conditional statement "if *p* then *q*" and the assertion not-*q*: they might suppose *p*, conclude *q* by MP, note that this contradicts the assertion, and conclude that their supposition of *p* was incorrect. This is a somewhat long chain of reasoning, but that very fact might account for all of the results: if this is the mechanism typically used, we would expect that neither mathematicians nor students would endorse all straightforward MT inferences by simple recognition, but that experienced mathematicians and more successful students would be better able to reach correct conclusions by correctly reasoning through the whole chain.

Prior to such investigations, we do not suggest that we should stop teaching mathematics students the material interpretation of the conditional. However, we do suggest that we should not be too concerned if undergraduate students do not develop to a point at which they reliably endorse MT inferences, because it appears that they may not need to.

References

Attridge, N., & Inglis, M. (2013). Advanced mathematical study and the development of conditional reasoning skills. *PLoS ONE*, *8*, e69399.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics*, *53*, 5-34.

Epp, S. (2003). The role of logic in teaching proof. *American Mathematical Monthly*, *110*, 886-899.

Evans, J. St. B. T., Clibbens, J., & Rood, B. (1995). Bias in conditional inference: Implications for mental models and mental logic. *Quarterly Journal of Experimental Psychology*, *48A*, 644-670.

Hoyles, C., & Küchemann, D. (2002). Students' understanding of logical implication. *Educational Studies in Mathematics*, *51*, 193-223.

Inglis, M., & Simpson, A. (2006). The role of mathematical context in evaluating conditional statements. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková, (eds.), *Proceedings of the 30th International Conference on the Psychology of Mathematics Education* (Vol. 3, pp. 337-344). Prague, Czech Republic. IGPME.

Inglis, M., & Simpson, A. (2009). The defective and material conditionals in mathematics: Does it matter? In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis, (eds.), *Proceedings of the 33rd International Conference on the Psychology of Mathematics Education* (Vol. 3, pp.225-232), Thessaloniki, Greece: IGPME.

Mejía -Ramos, J.-P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*, 3-18.

NCTM (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Quine, W. V. O. (1966). *Methods of Logic (second edition)*. Routledge & Kegan Paul: London.

A COMPARISON OF FOUR PEDAGOGICAL STRATEGIES IN CALCULUS

Spencer Bagley
San Diego State University and UC San Diego

The quality of education in introductory calculus classes is an issue of particular educational and economic importance. In work related to a national study of college calculus programs conducted by the MAA, I report on a study of four different pedagogical approaches to Calculus I at a single institution in the Fall 2012 semester. Using statistical methods, I analyze the effects of these four approaches on students' persistence in STEM major tracks, attitudes and beliefs about mathematics, and procedural and conceptual achievement in calculus. Using qualitative methods, I draw links from the statistical results to differences and commonalities in the four classroom strategies.

Key words: Calculus; Persistence; Achievement; Affect, Beliefs, and Attitudes; Classroom Research

The quality of education in science, technology, engineering, and mathematics (STEM) fields is an issue of particular educational and economic importance. The President's Council of Advisors on Science and Technology (PCAST, 2012) recently reported that fewer than 40% of students who originally intend to major in a STEM field actually complete a STEM degree; many of the reasons students give for switching implicate uninspiring, dull, or unimaginative introductory courses, or courses that are "over-stuffed" and taught too quickly (Seymour, 2006; Seymour & Hewett, 1997). A national study, funded by the NSF and under the auspices of the MAA, is currently being conducted examining the characteristics of successful programs in college calculus (CSPCC, 2012; Bressoud, Carlson, Mesa, & Rasmussen, 2013). Using the same surveys as the CSPCC study, my work builds on this project by examining the effects on student outcomes (including conceptual and procedural achievement, persistence in STEM majors, and beliefs and attitudes about mathematics) of different teaching strategies used at a single institution. Because of my focus on a single institution, rather than the hundreds studied by CSPCC, I also was able to collect longitudinal student enrollment data, detailed classroom observations, and measures of students' mathematical understanding.

Theoretical Perspective and Literature

The theoretical perspective undergirding this work is the emergent perspective (Cobb & Yackel, 1996), which holds that participation in classroom activity "constitute[s] the conditions for the possibility of learning" (p. 185), and that an individual's psychological development is enabled and constrained by their participation in classroom activities. Therefore, classes that present more opportunities for students' engagement and participation in classroom activities are seen as presenting more opportunities for robust student learning and improvement in beliefs and dispositions.

The development of productive attitudes, beliefs, and dispositions about mathematics is an important component of student success in calculus. Students' mathematical beliefs, including confidence, self-efficacy, and self-concept, correspond strongly with achievement in mathematics classes (Pajares & Miller, 1995; Carlson, 1999, Schommer-Aikins, Duell, & Hutter, 2005), as well as problem-solving behaviors (Carlson & Bloom, 2005). Further, students' beliefs and attitudes appear to be strongly influenced by the beliefs and teaching styles of their instructors (Schoenfeld, 1992).

One of the four classes under examination is an "inverted classroom" (Lage, Platt, & Treglia, 2000), in which lecture content is delivered outside of class time via internet videos, and class time is used by students to solve problems in small groups. Inverted classrooms have been studied in STEM fields including economics (Lage et al., 2000), physics (Deslauriers, Schelew, & Wieman, 2011), computer science (Gannod, 2007), and biology (Moravec, Williams, Aguilar-Roca, & O'Dowd, 2010). Save for a few recent reports in conference proceedings (e.g., Bowers & Zazkis, 2012; Overmyer, 2013; Wasserman, Norris, & Carr, 2013), there is a relative dearth of literature on the inverted classroom in undergraduate mathematics education.

Setting and Questions

In the Fall 2012 semester at a large public university in the southwestern United States, Calculus I was taught by four different instructors using four different instructional techniques. The first class was a traditional, lecture-based approach to calculus, together with recitation sections led by a TA. The second class was a more interactive, student-centered lecture without TA recitations. A third class was taught using an inverted model, as described above. The inverted class had no TA recitations, but the TAs were involved in helping answer student questions during class time. The final class was an interactive, student-centered, technology-intensive lecture with TA recitations, using applets developed in Geometer's Sketchpad to help develop students' intuition for calculus concepts. For ease of reference, I call these classes the traditional class, the interactive class, the inverted class, and the technology class, respectively.

This report is part of a broader study examining the following research questions:

RQ1: How do students in the four classes compare in their: a) persistence in STEM major tracks? b) attitudes, dispositions, and beliefs about mathematics? and c) conceptual and procedural achievement in calculus?

RQ2: How do students in the four classes compare to students in the CSPCC study, and specifically to students in successful programs, in their: a) persistence in STEM major tracks? and b) attitudes, dispositions, and beliefs about mathematics?

RQ3: How do the similarities and differences in opportunities for learning between the four classes contribute to the similarities and differences in outcomes?

This report focuses on RQ1, examining the effect of the different classes on student outcomes, and RQ3, identifying similarities and differences between the opportunities for learning presented by each class and linking them to the statistical results.

Data and Methods

This is a mixed-methods study whose data is both quantitative and qualitative. The quantitative data comprises student survey responses and scores on several assessments. The surveys, designed for use in the ongoing CSPCC study, were given at the start and end of the term (STS and ETS, respectively), and included questions about students' demographic information, mathematical preparation, beliefs and attitudes about mathematics, college and career plans, and their experience in Calculus I. I also collected student scores on the Calculus Concept Readiness (CCR) instrument (Carlson, Madison, & West, 2010), the Calculus Concept Inventory (CCI; Epstein, 2006), and a common final exam developed jointly by the four instructors.

I also collected qualitative data, including classroom observations and student focus group interviews. I observed and took extensive field notes in each of the class sessions that addressed related rates and the fundamental theorem of calculus. I also conducted focus group interviews with groups of four to seven student volunteers, focusing on students' subjective assessment of their calculus class. Both the interviews and the observations were audiorecorded and transcribed to aid in qualitative analysis, which was conducted using grounded theory (Strauss & Corbin, 1994).

To characterize and analyze the quantitative data, I used descriptive and inferential statistical methods. Tests such as ANOVA, ANCOVA, t-tests, and factor analysis were used to identify and assess the significance of differences between student outcomes in the four classes.

Results

In this section, I will summarize selected quantitative and qualitative results. The quantitative results fall into four broad categories: persistence, beliefs and attitudes, procedural and conceptual achievement, and differential impact on various subgroups.

Persistence

I identified *switchers* and *persisters* following the methodology used in the MAA study (Rasmussen & Ellis, 2013). Persisters are those who both began and ended the term intending to take Calculus II; switchers are those who initially intended to take Calculus II but changed their minds during the course of the term. I identified a total of 22 switchers and 253 persisters in the four classes. I conducted a chi-square analysis to determine if there were significant differences between the proportions of switchers and persisters in the four classes. The test reported no significant difference (p = .239); it thus appears that the differences in the four classes have no impact on students' persistence.

Table 1: Switcher and Persister Counts

		Tech.	**Interactive**	**Trad.**	**Inverted**	**Total**
Switchers	Count	9	0	8	5	22
	Expected	6.2	2.6	8.2	5.1	22.0
Persisters	Count	68	32	94	59	253
	Expected	70.8	29.4	93.8	58.9	253.0

Beliefs and Attitudes

There were 16 beliefs items on the ETS; I used ANOVA to determine if there were significant differences between the classes in the responses to the ETS items. Only five items were identified as differing significantly between classes. Post-hoc comparisons were conducted to determine the precise location of differences. The results are summarized in the table below; unless otherwise indicated, higher scores indicate more favorable beliefs.

Table 2: ANOVA ETS Beliefs Items

Item	$F(3, 325)$	p	Differences	MD	p
This course has increased my interest in taking more mathematics.	8.666	< .001	Tech. > Inverted Intr. > Trad. Intr. > Inverted	.688 .849 1.259	.007 .003 < .001
I am good at computing derivatives and integrals.	4.953	.002	Tech. > Inverted Trad. > Inverted	.564 .547	.008 .006
I am able to use ideas of calculus to solve word problems that I have not seen before.	2.718	.045	Trad. > Inverted	.502	.038
* My score on my mathematics exam is a measure of how well: (1 = I understand the covered material; 4 = I can do things the way the teacher wants)	3.766	.011	Tech. < Inverted Trad. < Inverted	.410 .392	.021 .018
When studying mathematics in a	4.235**	.007	Trad. > Tech.	.373	.030†

textbook, I tend to: (1 = memorize it the way it is presented; 4 = make sense of the material so that I understand it)			Trad. > Inverted	.413	.044†

* Reverse-coded; ** Welch $F(3, 137.994)$; † Tamhane's T2

As a dimension-reduction technique, I employed principal component analysis separately on the STS and ETS beliefs items. High KMO values (.808 and .848, respectively) and significant results from Bartlett's test of sphericity indicated that the correlation matrices should be factorable. On the STS, I retained two components, which explained 25.6% and 9.5% of variance, respectively, for a total of 35.1%. Two components from the ETS items were also retained, explaining 29.6% and 9.6% of variance, respectively, for a total of 39.3%.

After theoretical interpretation, I found that the first factor on both the STS and ETS measured students' affective beliefs about mathematics. ANOVA revealed no significant difference between the classes on the ETS affective beliefs variable. I used ANCOVA to compare scores on the ETS affective beliefs variable while controlling for scores on the STS affective beliefs variable, but again, no significant differences were revealed.

Procedural and Conceptual Achievement

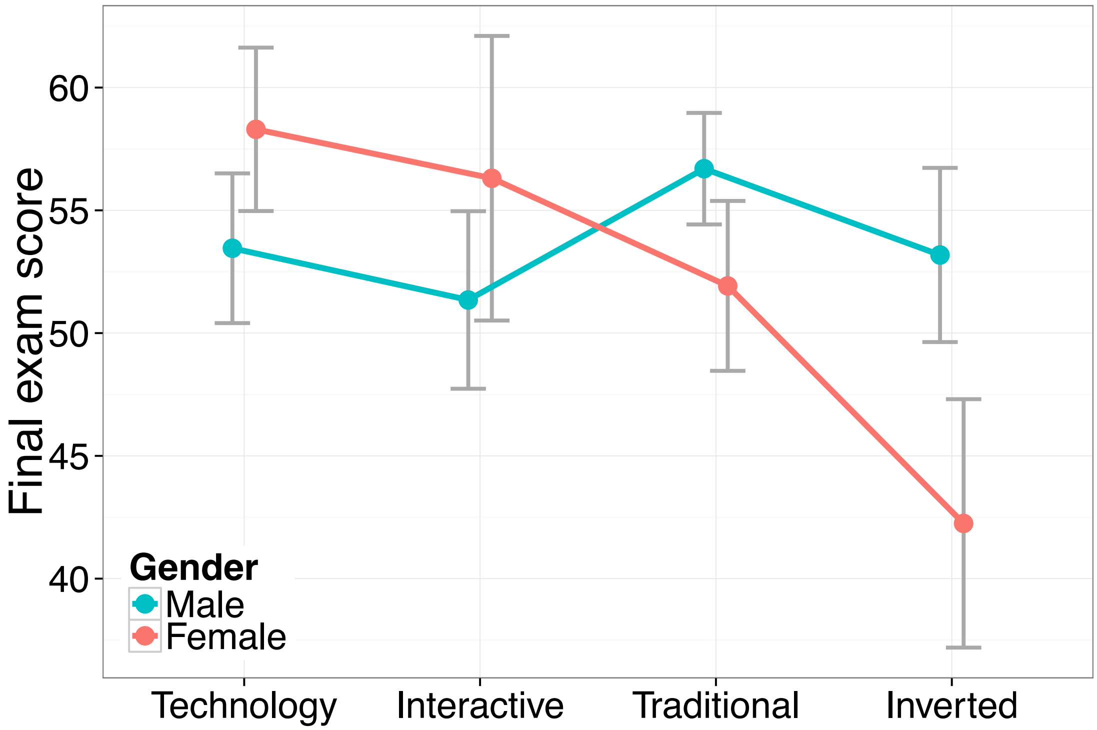
As a measure of growth in student understanding of calculus concepts, normalized gain on the CCI was computed by taking the ratio of actual gain (post-term – pre-term) to possible gain (maximum – pre-term) (Hake, 1998). ANOVA revealed no significant differences between the four classes in average normalized gain. Similarly, there were no significant differences between the four classes in post-term CCI score.

ANOVA revealed significant differences between the classes in raw percentage scores on the common final exam ($F(3, 429) = 5.145$, $p = .002$). Post-hoc comparisons using Tukey's HSD test showed that inverted-class students were outperformed by both technology-class students (MD = 7.09, $p = .032$) and traditional-class students (MD = 7.58, $p = .008$), and that interactive-class students outperformed traditional-class students (MD = 7.25, $p = .047$). However, when using ANCOVA to control for preparation as measured by the CCR instrument, no significant differences were detected between the classes on mean final exam score. The overall mean final exam score was 51.7% with a median of 53%.

Differential Impact

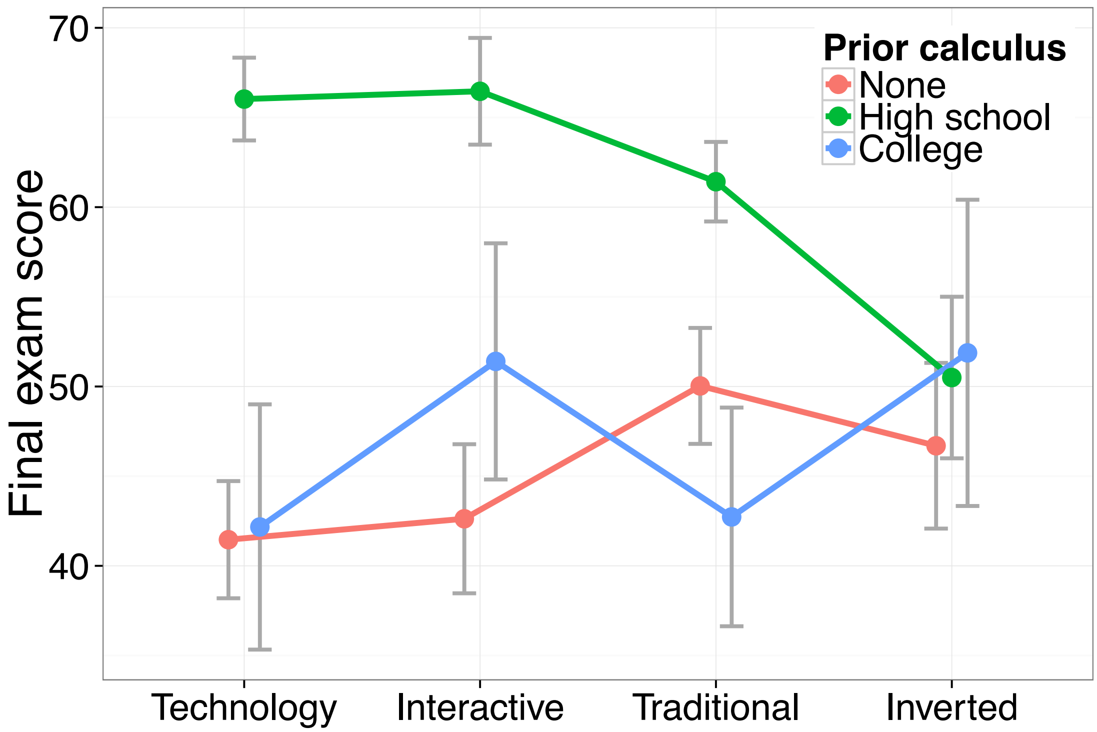
I conducted a number of two-way ANOVAs to assess the differential impact of the four classes on various different populations. For instance, I compared the performance of males and females on the final exam in each of the four classes. The two-way ANOVA showed no statistically significant main effect of either gender or class; the interaction approached significance ($F(3, 236) = 1.97$, $p = .119$). A marginal means plot is displayed in Figure 3. Univariate contrasts revealed that, as suggested by the marginal means plot, males and females did not differ statistically significantly in any of the classes but the inverted class, in which males outperformed females by 10.93 points ($t(36.97) = 1.77$, $p = .04$).

Figure 3: Comparison of Final Exam Scores by Gender



I also categorized students by their earliest prior calculus experience: high school (either AP or non-AP), college, or none. Two-way ANOVA revealed a significant main effect of prior calculus experience ($F(2, 233) = 20.67$, $p < .001$) and a significant interaction between prior calculus experience and class ($F(6, 233) = 2.63$, $p = .02$). The marginal means plot in Figure 4 suggests that students who took calculus in high school outperformed all other students in every class but the inverted class; indeed, it is only in this class that the univariate contrast is not significant.

Figure 4: Comparison of Final Exam Scores by Prior Calculus Experience



Qualitative Results

Given how different the classes seem, there are surprisingly few statistically reliable differences between student outcomes. It is natural, then, to inquire about the similarities between the classes. The qualitative data indicate that students' perceptions of the four classes share several important commonalities. One prominent theme that emerged from focus group interviews was that calculus is a prerequisite for calculus (triangulating the quantitative result discussed earlier). For instance, one student felt that the instructor of the technology class "teaches the class for people who have already taken calculus." A student in the inverted class said that students with no prior calculus are put at "a high disadvantage;" another said he felt bad for people in the class who hadn't taken calculus before. A student in the traditional class counted himself "lucky" to have taken calculus before, and said that "[when] I put myself in the shoes for people who are just learning it for the first time, I was like wow, really tough."

Another common theme from the focus group interviews was that students had concerns about the pacing of each day's class sessions. One student in the traditional class said that the instructor would often skip steps when presenting examples on the board, and he would wonder, "How did she get from this line to that line? … Can I see the step-by-step, please? But I think it's just time maybe. She's trying to cram everything." Similarly, one student in the interactive class felt that she did not have enough time to formulate questions when the instructor would ask the class if they understood something: "I need just a few more seconds to get what he's saying… and then I'll look up, and he already started doing a new example or something." However, she continued, "I get why, because we only meet twice a week." The fast pace of daily classes, which concerned the students in the focus groups, was probably linked to their instructors' felt need to cover a particular curriculum.

The focus group students in the inverted class were uniformly and vociferously dissatisfied with the implementation of the inverted model. The instructor did not make the videos himself, instead choosing them from online resources such as Khan Academy; the general feeling of the students was thus that “the problems on the [in-class] worksheet have no relation to the videos,” or that “the videos are not applicable to the work.” This had negative impacts on students’ confidence: “I watched the videos and I understand it going in, I feel very confident, and then I get that paper [the in-class worksheet] and I'm like, well, I give up already.” Another student said that she was “nervous for Calc II.”

Further, the instructor did not come to the daily class sessions, sending TAs in his stead. Students were dissatisfied with this practice. One student noted that there are questions that TAs cannot answer, no matter how smart the TAs are: “when we ask the TAs, what does [the instructor] want with this problem, they say ‘I don’t know.’” When asked what he thought was the instructor’s attitude toward students, another student replied, “I feel like we’re kind of a nuisance to him.” One student summed up the general feeling by saying: “I think the entire class is just kind of fed up with the whole thing.” Another said that the inverted model is “a good idea, but I don’t think it was put into practice very well.” Data from his course evaluations triangulated this finding: of the 36 students who responded to the open-ended comment prompt, 33 left comments negatively evaluating the inverted model.

Discussion and Future Directions

One obvious similarity between the classes is the shared curriculum mandated by the common final. Given the poor performance on the common final exam (recall that the overall mean was 51.7%), these results appear to confirm Seymour’s (2006) previous findings that introductory courses including Calculus I are often “over-stuffed” and taught too quickly. My data appear to be one more piece of evidence supporting the ongoing push for a “lean and lively calculus” (see, e.g., Steen, 1988). Many studies over the past thirty years have supported this conclusion; perhaps we in the mathematics education community need to find new ways to communicate to administrators and instructional designers that the current curriculum is “too much, too fast” for students to master. It is worth observing that one of the instructors in this study is a mathematics education researcher with a great deal of theoretical and practical knowledge. The fact that this instructor could not be distinguished statistically from the others, even with wealth of theoretical and practical knowledge that informs their teaching, is illustrative of the constraints imposed by the curriculum.

My data on students’ reception of this implementation of the inverted model support the literature’s growing consensus on a set of best practices for inverting a classroom. While searching the literature on the inverted model, I identified a set of commonalities among the most successful inverted classrooms: first, pre-lecture activities must be made by the instructor; second, students must be held accountable for completing the pre-lecture activities; and third, time vacated by lecture must be replaced with active-learning exercises with the full participation of the instructor. It appears that students in my study were dissatisfied with this implementation of the inverted model precisely because it did not include these best practices.

Future analysis planned in this research program includes comparisons, both statistical and qualitative, of these four calculus classes with other calculus classes nationwide, and particularly those at institutions identified by the MAA study as particularly successful. Additionally, I plan to use multivariate regression to build a profile of the typical “switcher,” or student who chooses not to persist in a STEM major track, and compare the profile produced by my data with the profile produced by the MAA study’s nationwide data. In my conference talk, I will present relevant portions of these analyses, as well as more detailed segments of the quantitative and qualitative results reported here.

References

Bowers, J., & Zazkis, D. (2012). Do students flip over the "flipped classroom" model for learning college calculus? In L. R. Van Zoest, J.-J. Lo, & Kratky, J. L. (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (849-852). Kalamazoo, MI: Western Michigan University.

Bressoud, D., Carlson, M., Mesa, V., & Rasmussen, C. (2013). The calculus student: Insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology*, *44*(5), 685-698. doi:10.1080/0020739X.2013.798874

Carlson, M. (1999). The mathematical behavior of six successful mathematics graduate students: Influences leading to mathematical success. *Educational Studies in Mathematics*, *40*(3), 237–258.

Carlson, M., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics*, *58*(1), 45–75. doi:10.1007/s10649-005-0808-x

Carlson, M., Madison, B., & West, R. (2010). The Calculus Concept Readiness (CCR) Instrument: Assessing Student Readiness for Calculus. *arXiv preprint arXiv:1010.2719*. Retrieved from http://arxiv.org/abs/1010.2719

Characteristics of Successful Programs in College Calculus. (2012). *Characteristics of successful programs in college calculus.* Retrieved from http://www.maa.org/cspcc/

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, *31*(3-4), 175–190. doi:10.1080/00461520.1996.9653265

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, *332*(6031), 862–864. doi:10.1126/science.1201783

Epstein, J. (2006). The calculus concept inventory. *National STEM Assessment, Washington, DC*, 60-67.

Gannod, G. C., Burge, J. E., & Helmick, M. T. (2008). Using the inverted classroom to teach software engineering. *Proceedings of the 30th international conference on Software engineering* (pp. 777–786). Retrieved from http://dl.acm.org/citation.cfm?id=1368198

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*, 64.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1), 30. doi:10.2307/1183338

Moravec, M., Williams, A., Aguilar-Roca, N., & O'Dowd, D. K. (2010). Learn before Lecture: A Strategy That Improves Learning Outcomes in a Large Introductory Biology Class. *Cell Biology Education*, *9*(4), 473–481. doi:10.1187/cbe.10-04-0063

Overmyer, J. (2013, February). *The flipped classroom model for college algebra: Effects on student achievement.* Poster presented at the Conference on Research in Undergraduate Mathematics Education, Denver, CO.

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of counseling psychology*, *42*(2), 190.

President's Council of Advisors on Science and Technology (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Washington, D.C.: Office of Science and Technology Policy.

Rasmussen, C., & Ellis, J. (2013). Who is switching out of calculus and why? In Lindmeier, A. M. & Heinze, A. (Eds.). *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education, Vol. 4* (pp. 73-80). Kiel, Germany: PME.

Rasmussen, C., Ellis, J., Duncan, K., Bressoud, D., & Carlson, M. (2013). Who are the students who switch out of calculus and why? Paper presented at the Annual Meeting of the American Educational Research Associations, San Francisco, CA.

Rasmussen, C., Yackel, E., & King, K. (2003). Social and sociomathematical norms in the mathematics classroom. In H. Schoen, & R. Charles (Eds.), *Teaching mathematics through problem solving: Grades 6–12* (pp. 143–154). Reston, VA: NCTM.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334–370). New York: MacMillan.

Schommer-Aikins, M., Duell, O. K., & Hutter, R. (2005). Epistemological beliefs, mathematical problem-solving beliefs, and academic performance of middle school students. *The Elementary School Journal*, *105*(3), 289–304. doi:10.1086/428745

Seymour, E. (2006). *Testimony offered to the Research Subcommittee of the Committee on Science of the U.S. House of Representatives hearing on Undergraduate Science, Mathematics, and Engineering Education: What's Working?* Available from http://commdocs.house.gov/committees/science/hsy26481.000/hsy26481_0f.htm

Seymour, E., & Hewett, N. (1997). Talking about leaving: why undergraduates leave the sciences. Boulder, Colo: Westview Press.

Steen, L. A. (1988). *Calculus for a New Century: A Pump, Not a Filter. MAA Notes Number 8*. Washington, DC: MAA.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of qualitative research*, 273-285.

Wasserman, N., Norris, S., & Carr, T. (2013, February.) *Comparing a "flipped" instructional model in an undergraduate Calculus III course.* Paper presented at the Conference on Research in Undergraduate Mathematics Education, Denver, CO.

STUDENT UNDERSTANDING OF THE FUNDAMENTAL THEOREM OF CALCULUS AT THE MATHEMATICS-PHYSICS INTERFACE

Rabindra R. Bajracharya, John R. Thompson
University of Maine

We studied students' understanding of the Fundamental Theorem of Calculus (FTC) in graphical representations that are relevant in physics contexts. Two versions of written surveys, one in mathematics and one in physics, were administered in multivariable calculus and introductory calculus-based physics classes, respectively. Individual interviews were conducted with students from the survey population. A series of FTC-based physics questions were asked during the interviews. The written and interview data have yielded evidence of several student difficulties in interpreting or applying the FTC to the problems given, including attempting to evaluate the antiderivative at individual points and using the slope rather than the area to determine the integral. The interview results further suggest that students often fail to make meaningful connections between individual elements of the FTC.

Key words: Fundamental Theorem of Calculus, Physics, Difficulties, Representations

Introduction

We have been exploring the effect of student understanding of various concepts in mathematics on their understanding of physics concepts and vice versa. Learning physics concepts often requires the ability to interpret and manipulate the underlying mathematical representations (e.g., equations, graphs, and diagrams). A proper understanding of representations of physics concepts often requires identification of the relationship between the physics and the mathematics built into the representation as well as subsequent application of the mathematical concepts (Chi et al., 1981; Redish, 2005). Several studies in physics education research (PER) indicate connections between students' understanding of mathematics concepts and their understanding of physics concepts. Some PER findings suggest that some of the student difficulties categorized as physics difficulties may be related to the mathematics and its representations in addition to, or instead of, being difficulties with the physics (Christensen & Thompson, 2010; Christensen & Thompson, 2012; Meltzer, 2002; Pollock et al., 2007).

One interesting aspect of student understanding is the ability to relate mathematical concepts learned in a mathematics class to various physics concepts. One topic that plays a significant role in physics is the Fundamental Theorem of Calculus (FTC), which is relevant in determining various physical quantities such as displacement, potential difference, and work. In order to fully understand the FTC, a working understanding of many concepts, such as function, rate of change, antiderivative, definite integral, is needed. Research in undergraduate mathematics education attributes student difficulty with the FTC primarily to students' difficulty with the function concept (Carlson et al., 2003; Thompson, 1994; Thompson, 2008) and rates of change (Thompson, 1994).

Connecting student understanding of mathematics and physics is relevant to mathematics educators as well, since many mathematics courses use various basic physics topics for applications of mathematics concepts. In calculus, topics such as displacement, velocity and mechanical work are used as contexts for understanding integrals and derivatives. Studies have shown students using physics concepts while attempting to understand or interpret mathematical concepts (Bajracharya et al., 2012; Marrongelle, 2004). In fact, researchers have suggested the use of physical contexts (e.g. displacement, velocity, etc.) when introducing the FTC (Rosenthal, 1992; Schnepp & Nemirovsky, 2001). However, it may be that students who are unable to understand the physics concepts in the applied context may have more difficulty understanding the mathematical concepts being taught.

Similarly, physics students are often expected to be able to find connections between the rate of change (derivative) and the accumulation (definite integral) of a physical quantity (function), particularly based on graphical representations. However, to our knowledge there is no explicit research on student understanding of FTC concepts in physics, despite its ubiquitous use in various physics contexts. Researchers in physics education have studied student interpretation and use of graphs in kinematics. Beichner (1994) found that students did not recognize the physical meaning of areas under kinematics graph curves, and that students often performed slope calculations or subtracted axis values when an area calculation was required, regardless of what was graphed. We are exploring the extent to which students' understanding of the FTC affects their basic physics problem solving. Being able to distinguish whether students are struggling with the physics ideas or the underlying mathematics (or both) can inform instruction in both disciplines to help students connect the mathematics and the disciplinary contexts in which that mathematics is applied.

Theoretical Perspective

We have been investigating the conceptual difficulties that students have with the FTC in graphical representations using the notion of *specific student difficulties* (Heron, 2003). According to this perspective, students manifest their difficulties through incorrect or inappropriate ideas, or flawed patterns of reasoning to specific questions. Identification of specific student difficulties is a pragmatic approach that has led to the development of research-validated instructional strategies and materials that have improved students' conceptual understanding in many contexts across the physics curriculum (e.g., McDermott, 2001). Specific difficulties are typically identified through empirical studies and are crucial for building theoretical models of student thinking because they could be used to verify those models.

The specific difficulties perspective does not necessarily speak to the origins of the difficulties being identified. There are other theoretical frameworks common in physics that address this to varying extents (e.g., resources, misconceptions). Often a difficulty can be due to the inappropriate application of a reasonable idea, an undeveloped distinction between two related concepts (e.g., velocity and acceleration (Trowbridge and McDermott, 1981)), or an incorrect naïve "theory" (e.g., a conception of *impetus* in a moving body associated with net force (Clement, 1982).

Methodology

We have constructed questions, often with parallel versions in both mathematics and physics, that either explicitly or implicitly requires the application of the FTC in a graphical interpretation. These questions were administered as written surveys in lecture sections of second-semester calculus-based introductory physics and multivariable calculus for two consecutive semesters. A total of 159 mathematics and 90 physics students participated during the first survey. During the second survey administration, 92 mathematics and 120 physics students participated. Here we focus on only one pair of questions (Fig. 1).

We also conducted 13 individual interviews to probe the depth and breadth of students' understanding and application of the FTC in physics that were not revealed in the survey results, as well as the robustness of the explanations and lines of reasoning seen in the written responses. Subjects were asked five FTC-based questions in physics contexts of varying familiarity. However, these questions could be answered using the FTC without any prior knowledge of the physics. (Figure 2 depicts an example question.) The solutions to the first two questions required explicit use of the given graphs (i.e., determination of the area under the curve between the integration limits). The next two could be solved either graphically or analytically, using a given algebraic function. The last one required a numerical solution.

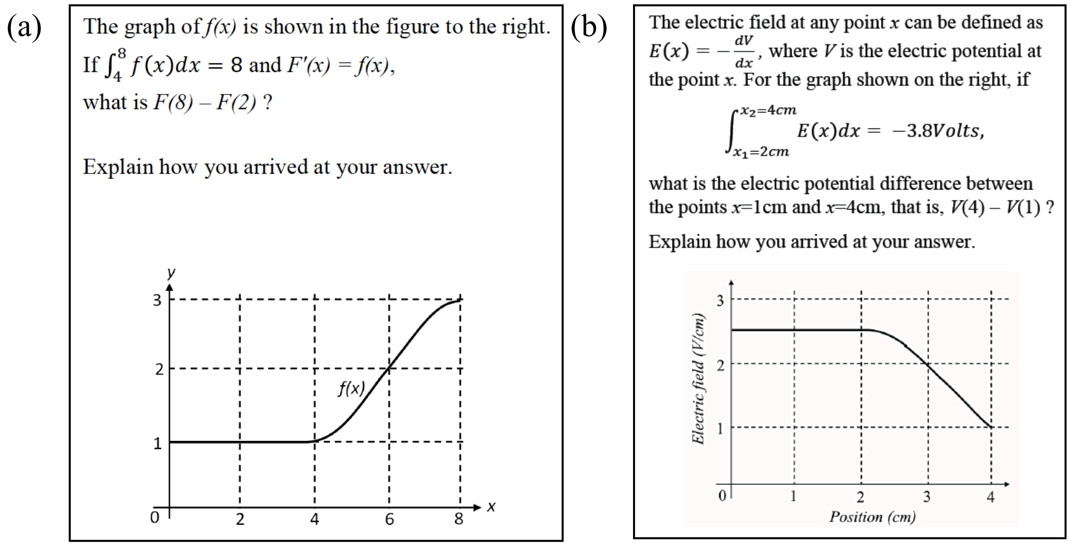


FIGURE 1. Analogous (a) mathematics and (b) physics versions of the written surveys.

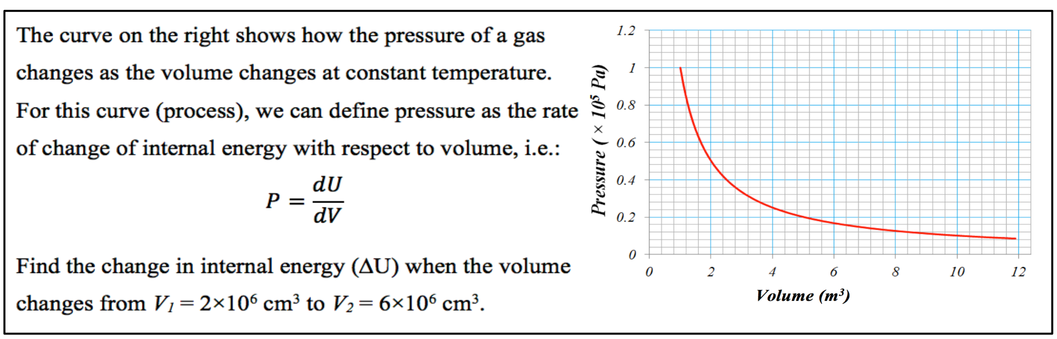


FIGURE 2. An interview question requiring explicit use of the graph.

Data analysis used grounded theory (Strauss & Corbin, 1997) but paid attention to specific difficulties related to the FTC.

Results

In the written surveys, about half of the students in both the mathematics and physics classes gave correct responses. Students used various reasoning strategies to answer the questions. Five strategies are described below; four of these indicate student difficulties with the FTC.

1. Connecting the integral, antiderivative, and area under the curve. (Fig. 3.) Most students who provided correct responses used the FTC explicitly or implicitly. These students equated $F(b) - F(a)$, the area under the curve, and the definite integral ($\int_a^b f(x)dx$), effectively counting the squares under the curve between the limits to find the desired quantity. While most students used the correct area, a few (<5%) chose the base for their area calculations as the horizontal line that passes through the endpoint of the curve (e.g., y=1 in Fig. 1(a)) rather than the x-axis.

2. Evaluating individual antiderivative values at endpoints. (Fig. 4.) One group of students evaluated the individual values of the antiderivatives at endpoints (e.g., $F(b)$ and $F(a)$). Finding the individual antiderivatives leads to a correct answer when they consider each of them to be equal to the areas under the curve between a common lower limit (here $F(0)$) and the upper limits as shown in Fig. 4a. However, this was not a consistently correct approach, as students also used other computational approaches to find the individual antiderivatives, as in Fig. 4b. This suggests difficulty recognizing that the difference in antiderivative values at the endpoints (e.g., $F(b) - F(a)$) is the definite integral of the given function between the given limits, and is related to the area under the curve in the given interval.

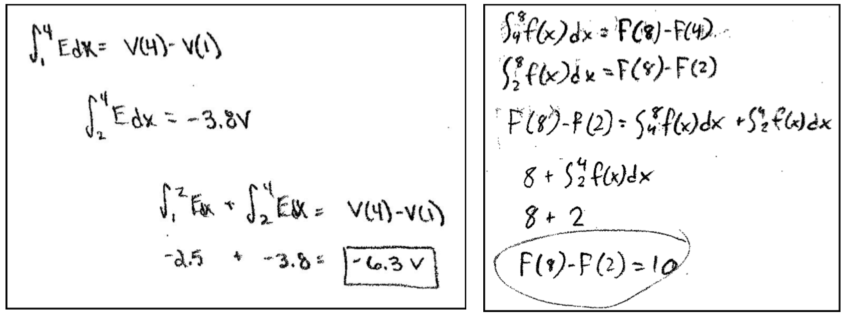


FIGURE 3. Students connecting the integral, antiderivative, and area under the curve.

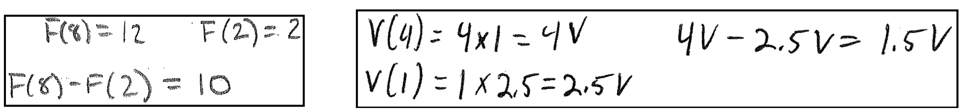


FIGURE 4. Students evaluating individual antiderivative values at end points.

3. Confusing antiderivative and function. (Fig. 5.) One of the most common responses was to use the difference of the original function at the endpoints (i.e. $f(b) - f(c)$) rather than the difference of the antiderivative at the endpoints (i.e. $F(b) - F(c)$), suggesting an operational confusion between the antiderivative and the function in a graphical context. This is consistent with earlier findings in upper-division thermodynamics courses in which students used the difference of endpoint values to compare the works done on a system during two different thermodynamic processes (Pollock et al., 2007).

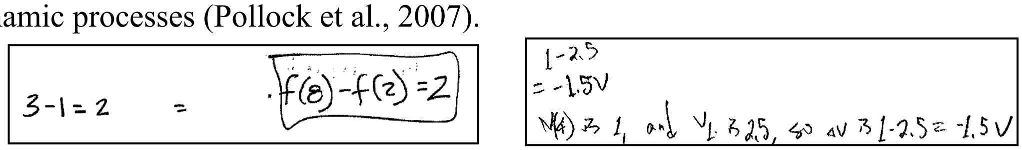


FIGURE 5. Students confusing antiderivative and function.

4. Confusing slope or derivative with area. (Fig. 6.) A few students provided their responses using slope-based computational reasoning. Some students evaluated the slope over the interval (i.e., $\Delta y/\Delta x$) as the required answer, whereas others tried different slope-based properties, such as $F(1) = F(0)$, in their responses.

5. Reasoning analytically. (Fig. 7.) Students in this category approached the problems in two distinct ways: approximating the given curve with an algebraic function, inserting that function as the integrand, and integrating; or

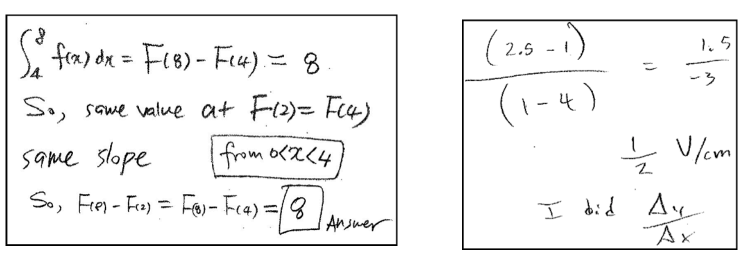


FIGURE 6. Students confusing slope or derivative with area.

considering the given numerical value of the integral as a function. We do not know the extent to which they understand the FTC, since their computations do not reflect relevant operations in these problems. Previous studies have also documented students' difficulties with problems without algebraic functions (Selden et al., 1989; Selden et al., 2000). This solution type is also consistent with the *action* view of function (Dubinsky & Harel, 1992; Oehrtman et al., 2008).

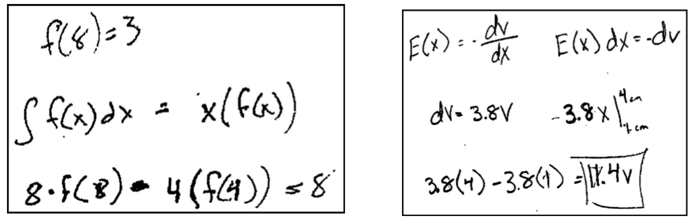


FIGURE 7. Students using analytical reasoning.

Interview results generally supported written data. The majority of students failed to use the FTC to determine the physical quantities, e.g., the change in internal energy, when the question did not include an algebraic function explicitly. For questions explicitly involving functions, most students took the antiderivative right away and solved the problem correctly. When subsequently prompted to answer these questions using a different approach, they concluded that the solution could be represented by the area under the curve. Our interview results suggest that even students who have a sufficient understanding of each of the constituents of the FTC (e.g., function, rate, area) often fail to see the connection between these elements. We speculate that their inability to find the right connections between the elements of the FTC results in the various kinds of previously mentioned specific difficulties.

Conclusions

Our preliminary results describe specific student difficulties with the Fundamental Theorem of Calculus common to both mathematics and physics contexts. Some of our findings in this research agree with previously reported difficulties (Beichner, 1994; Thompson, 1994). The interview results revealed that the majority of the physics students have a good grasp of most of the individual components of the FTC, but often fail to connect these components to solve problems. However, when dealing with unfamiliar physics contexts and without an analytical expression from which to start, either students struggle to meaningfully connect the individual elements of the FTC or their difficulties with even one element hinder their attempts to find the meaningful connection between the individual elements of the FTC. We are analyzing the interviews in greater depth to see where in the protocol students recognize the appropriate connections and the extent to which physics context affects their performance.

Our preliminary interview analyses indicate that students use different strategies to solve the FTC-based physics problems. Although attempted initially, we did not analyze our data using the notion of transfer as preparation for future learning (e.g., Schwartz et al., 2005) because the interviews we conducted did not fit the required format described by Schwartz et al. Our ongoing work includes interview analysis using the lens of *epistemic games*, which are defined as a set of rules and strategies that that are guided by a specific purpose, e.g. learning a concept (Collins & Ferguson, 1993). Our approach consists of comparing our grounded-theory-based problem-solving strategies to existing, identified epistemic games (Collins & Ferguson, 1993; Tuminaro and Redish, 2007).

Questions for audience

Which theoretical framework(s) or model(s) might be most appropriate to interpret students' difficulties connecting the elements of the FTC?

Are there specific aspects of the FTC and its related concepts that may need to be considered here but aren't?

Often the representations and notations used in mathematics and physics are very different. This often hinders students' ability to access ideas from mathematics to use in physics or vice versa. Is there a way to deal with this issue in both directions?

References

Bajracharya, R. R., Wemyss, T. M., & Thompson, J. R. (2012) Student interpretation of the signs of definite integrals using graphical representations, *AIP Conf. Proc.* 1413, 111-114.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics, 62* (8), 750-762.

Carlson, M. P., Smith, N., & Persson, J. (2003). Developing and connecting calculus students' notation of rate of change and accumulation: The fundamental theorem of calculus. In N. A. Pateman, B. J. Doughety, & J. T. Zillioz (Ed.), *Proceedings of the 2003 Joint Meeting of PME and PMENA. 2*, pp. 165-172. Honolulu, HI: CRDG, College of Education, University of Hawai'i.

Chi, M. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5* (2), 121-152.

Christensen, W. M. and Thompson, J. R. (2012). Investigating graphical representations of slope and derivative without a physics context. *Phys. Rev. ST Physics Ed. Research* 8, 023101.

Christensen, W. M., & Thompson, J. R. (2010). Investigating student understanding of physics concepts and the underlying calculus concepts in thermodynamics. *Proceedings of the 13th Annual Conference on Research in Undergraduate Mathematics Education, MAA*.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics*, 50(1), 66-71

Collins, A., & Ferguson, W. (1993). Epistemic forms and games. *Educational Psychologist*, 28(1), 25–42

Dubinsky, E., Harel, G. (1992). The nature of the process conception of function. In G. Harel & E. Dubinsky (Eds.), *The concept of functions: Aspects of epistemology and pedagogy*, MAA Notes, 25, 85-106.

Heron, P. (2003). Empirical investigations of learning and teaching, Part I: examining and interpreting student thinking. *Proceedings of Enrico Fermi Summer School on Physics Education Research.* (pp. 341-350). Italian Physical Society.

Marrongelle, K. A. (2004). How students use physics to reason about calculus tasks. *School Science and Mathematics*, 104(6), 258-272.

McDermott L. C. (2001). Oersted Medal Lecture 2001: Physics Education Research –The Key to Student Learning. *American Journal of Physics*. *69*, 1127-1137.

McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics, 55* (6), 503-513.

McDermott, L. C., Shaffer, P. S., and the Physics Education Group at the University of Washington (2002). *Tutorials in Introductory Physics*, 1st ed. (Prentice Hall, 2002).

Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics, 70* (12), 1259-1268.

Oehrtman, M. C., Carlson, M. P., and Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. P. Carlson & C.

Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 27-42). Washington, DC: Mathematical Association of America.
Pollock, E. B., Thompson, J. R., and Mountcastle, D. B. (2007). Student understanding of the physics and mathematics of process variables in P-V diagrams. In L. Hsu (Ed.), *2006 Physics Education Research Conference* (pp. 168-171). Greensboro, NC: AIP Conference Proceedings.
Redish, E. F. (2005). Problem Solving and the Use of Math in Physics Courses. *Proceedings of the Conference: World View on Physics Education in 2005: Focusing on Change.* Delhi.
Rosenthal, B. (1992). Discovering and Experiencing the Fundamental Theorem of Calculus. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 2(2), 131-154.
Schnepp, M., & Nemirovsky, R. (2001). Constructing a Foundation for the Fundamental Theorem of Calculus. In A. A. Cuoco & F. Curcio (Eds.), *The Roles of Representation in School Mathematics* (pp. 90-102). Reston, VA: NCTM.
Selden, J., Mason, A., and Selden, A. (1989). Can average calculus students solve non-routine problems? *Journal of Mathematical Behavior*, 8(1), 45-50.
Selden, A., Selden, J., Hauk, S., and Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? *Issues in Mathematics Education*, 8,128-153.
Strauss, A., & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park, AC: Sage.
Schwartz D. L., Bransford J. D., Sears D. (2005). Efficiency and innovation in transfer. In: Mestre J. P., editor. *Transfer of Learning from a Modern Multidisciplinary Perspective*. Greenwich, CT: Information Age Publishing, p. 1-51.
Trowbridge, D. E., & McDermott, L. C. (1981). Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics*, 49(3), 242–253.
Thompson, P. W. (1994). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics, 26*, 229-274.
Thompson, P. W. (2008). The concept of accumulation in calculus. In M. P. Carlson, & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp. 43-52). Washington, DC: Mathematical Association of America.
Tuminaro, J. and Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games, *Phys. Rev. ST Phys. Educ. Res.* 3, 020101.

TRANSFORMING REMEDIAL MATHEMATICS INSTRUCTION WITH HIGH-QUALITY PEER TEACHING

Kristen N. Bieda, Raven McCrory, Steven Wolf
Michigan State University

Background

In the United States, many students enter their first year of college unprepared to take college level mathematics: at public 4-year institutions, 16% of entering freshmen took remedial (developmental) mathematics in 2000 (U.S. DOE, 2003, p. 18). These students are less likely to enter or persist in STEM majors, and even less likely to graduate from college than students who are prepared for, and succeed in, college mathematics in their first year (Adelman, 2006). According to the recent AAU (2012) draft discussion document announcing their STEM initiative, about 25% of freshmen across the country intend to go into a STEM field, but only 15 -17% of graduates complete a STEM major (AAU, pp. 2-3). Most students who drop STEM majors do so during the first two years of college, often because of trouble in their first year mathematics courses (AAU, 2012, p. 4).

At least two issues are apparent here: First is the "preparation gap", the inadequate mathematical preparation of incoming college students. Institutions of higher education commonly offer developmental mathematics courses to address this problem, placing students in pre-college level algebra in an effort to get them ready to succeed in higher-level mathematics and science courses. The root cause of this problem, though, lies not in the postsecondary institution, but in the preparation of students before they reach college, in their elementary, middle and high schools. Many point to the inadequate preparation of mathematics teachers as an important factor, and call for better preparation of these teachers (Conference Board of the Mathematical Sciences, 2012; Schmidt, Blömeke, & Tatto, 2011). Although post-secondary institutions cannot directly influence what happens in high schools and earlier, they can work to improve what new K-12 teachers bring to K-12 mathematics classrooms.

The second problem is the retention of students in STEM majors once they enter college. Researchers and policy makers agree that one of the major problems in initial mathematics courses is inadequacy of pedagogy (e.g., AAU, 2011; Fairweather, 2008). Even though disciplinary-based researchers who study mathematics and science education have extensive evidence about how to teach introductory mathematics and science college courses in ways that "work", large scale adoption of effective teaching methods has been elusive. Faculty have been reluctant to change their teaching practices *even in the face of strong empirical evidence about what works* in part because the rewards for doing so are limited and it is hard work that requires substantial investment of time and effort (AAU, 2011).

This project investigated the potential of a hybrid remedial mathematics course (RMC), taught by a corps of undergraduate peers in a secondary mathematics teacher preparation program, to provide remedial mathematics students with opportunities to develop robust mathematical proficiency. Specifically, the question guiding our research is: *In a developmental mathematics course, what is the impact on students' mathematical proficiency of an intervention using teaching methods and materials designed to help them develop mathematical proficiency?* In the section that follows, we describe the theoretical framework guiding the instruction in this hybrid RMC.

Theoretical Framework

Research on effective mathematics teaching and learning shows that, although students learn in different ways, the goal of mathematical proficiency is best reached through methods that engage students in doing, talking, and thinking about mathematics. This is particularly the case for students in undergraduate remedial mathematics courses (Hodera, 2011). The National Research Council's report *Adding it Up,* authored by Kilpatrick, Swafford and Findell (2001), identified and defined five strands to explain what it means to be proficient in mathematics. The strands are defined as follows:

- *Conceptual understanding* is the "comprehension of mathematical concepts, operations, and relations"
- *Procedural fluency* is "knowledge of procedures, knowledge of when and how to use them appropriately, and skill in performing them flexibly, accurately, and efficiently."
- *Strategic competence* is the "ability to formulate, represent, and solve mathematical problems"
- *Adaptive reasoning* is the "capacity for logical thought, reflection, explanation, and justification"
- *Productive disposition* is the "habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy." (Kilpatrick, et al., p.116, 121)

Effective mathematics teaching attends to all five strands, aiming to help students learn to solve problems, apply mathematics in new contexts, and see mathematics as a coherent and powerful discipline. To achieve this, students should be engaged in problem solving and mathematical discourse, working in small groups and whole class discussions to explain mathematical ideas and justify mathematical solutions (Fuson, Kalchman & Bransford, 1999; Kilpatrick, Swafford & Findell, 2001). Research on classroom discourse in mathematics and on teaching through problem solving and discourse provides guidance about how to plan for and manage classes that maintain a rigorous mathematical focus while encouraging students to participate (Chapin, O'Connor & Anderson, 2003; Stein, Engle, Smith & Hughes, 2008). This project implemented a curriculum and pedagogy based on these principles -- students work collaboratively, talk about their reasoning, and complete tasks that build robust mathematical proficiency --- enacted by prospective secondary mathematics teachers (PSMTs).

Research Method

We piloted the project in Fall 2012 with 34 students in two sections of a RMC. Students in the RMC used an online adaptive tutoring program (ALEKS, Falmagne, Cosyn, Doignon, & Theiry, undated) to complete coursework and attended a face-to-face support class twice a week for two hours each. Ten PSMTs pairs rotated responsibility for instruction in each support class session. We collected achievement data from intervention and control sections of the RMC including placement (pretest) scores, final exam scores, and assessment from the ALEKS system. We collected video data from the intervention sections, and interviews with students from both intervention and control sections. We administered a pre- and post- attitudes and beliefs assessment based on Fennema (1976) and Bai, Wang, Pan, & Frey (2009). Demographic data for students in both intervention and control sections were obtained from the registrar. Analysis is ongoing and is using statistical methods that include comparisons of means, multilevel modeling, and, for the interviews and videos, discourse analysis.

Research Findings

We piloted the project in Fall 2012 with 34 students in two sections of a RMC each with 10 PSMTs pairs who rotated responsibility for instruction in each class session. While no significant differences (α = 0.05) were found, the students in the intervention had higher scores on average on both the final exam and the online system assessment than other students in the RMC. The Winsorized mean values for the intervention groups and two comparison groups are shown in Table 1. Because of a power issue, there is a 93% likelihood that a Type 2 error (failure to reject a false null hypothesis) occurred. Longitudinal data from our institution's mathematics department suggests that the difference in performance of the intervention students on the final exam does predict a higher grade in the subsequent mathematics course, College Algebra, by a half grade-point (e.g., 3.5 rather than 3.0). We are currently developing a multilevel regression model to control for factors such as prior knowledge, SES, and instructor.

Table 1
End-of-course results for intervention and comparison groups

		Final Exam (Max 200)		**Online Post Test (Max 248)**		**Online Gain**	
Group	**n**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	SD
Online students	701	134	51	209	45	149	46
Face-to-face Control	71	135	44	220	47	162	46
Face-to-face Intervention	**34**	**138**	**35**	**218**	**25**	**174**	**31**
All	806	134	50	211	44	151	46

The project, modified based on some of our findings from Year 1, will continue in 2013-14 with a second cohort of remedial mathematics students and PSMTs. Modifications include revisions to the course curriculum to better align with the needs of remedial mathematics students and condense the number of topics taught during the semester.

References

AAU (Association of American Universities, 2011). Five-year initiative for improving undergraduate STEM Education: Discussion draft, from http://www.aau.edu/policy/article.aspx?id=12588

Bai, H., Wang, L., Pan, W., & Frey, M. (2009). Measuring Mathematics Anxiety: Psychometric Analysis of a Bidimensional Affective Scale. [Article]. *Journal of Instructional Psychology, 36*(3), 185-193.

Chapin, S. H., O'Connor, M. C., & Anderson, N. C. (2003). *Classroom discussions: Using math talk to help students learn*. Sausalito, CA: Math Solutions Publications.

Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II: Draft for public discussion*. American Mathematical Society, Providence, RI.

Fairweather, J. (2008). *Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education*. The national Academies National Research Council Board of Science Education. Washington, D.C.

Falmagne, J.-C., Cosyn, E., Doignon, J.-P., & Theiry, N. (undated). *The Assessment of knowledge, in theory and in practice*. http://www.aleks.com/.

Fennema, E. (1976). Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*(5), 324-326.

Fuson, K. C., Kalchman, M., & Bransford, J. D. (1999). Mathematical understanding: An introduction. In J. D. Bransford, A. L. Brown & R. R. Cocking (Eds.), *How people learn: Brain, mind, experience, and school* (pp. 217-256). Washington, DC: National Academy of Sciences, National Research Council.

Hodara, M. (2011). Improving Pedagogy in the Developmental Mathematics Classroom. *CCRC Brief, 51* (pp. 1-4). Community College Research Center, Columbia University.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). Adding it up: Helping children learn mathematics. Washington, DC: National Research Council, Mathematics Learning Study Committee, National Academy Press.

Schmidt, W., Blömeke, S., & Tatto, M. T. (2011). *Teacher education matters: A study of middle school mathematics teacher preparation in six countries*. New York: Teachers College Press.

Stein, M.K., Engle, R.A., Smith, M.S. & Hughes, E.K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313-340.

U.S. Department of Education, National Center for Education Statistics. *Remedial education at degree-granting postsecondary institutions in fall 2000*, NCES 2004-010, by Basmat Parsad and Laurie Lewis. Project Officer: Bernard Greene. Washington, DC: 2003.

Preservice Teachers' Uses of the Internet to Investigate the Proof of the Pythagorean Theorem and its Converse

Aaron Brakoniecki
Michigan State University

Learners of mathematics, including preservice teachers, often explore online resources when investigating mathematical problems. When asked to search online for resources that would help them be able to better explain a proof of the Pythagorean theorem and its converse, preservice teachers used a variety of different searching strategies to locate information. Further, the ways in which this information was incorporated into their understanding of mathematics became evident through concept maps. This proposal describes the study conducted and initial results from the data and asks the reader to consider possible ways this research might be extended and refined.

Keywords: Preservice Teacher, Pythagorean Theorem, Internet

When learning about any mathematical content, many mathematics students utilize the internet as a resource to support their learning of mathematics. Preservice elementary teachers are no exception. Additionally, the mathematics content course that preservice elementary teachers are often required to take is limited when compared with the mathematical content of the grades these preservice teachers might be responsible for teaching once they begin their practice (K-5 or even K-8). The internet is likely a rich resource of information to support learning; however, it is not yet clear for what purposes these learners are going to online resources for, what strategies they use to locate information, and how they incorporate any information they find into their larger understanding of that topic. This study seeks to begin to unpack some of this mystery and shed light on how learners of mathematics (specifically preservice teachers) use the internet to support their learning of mathematics.

The Common Core State Standards for Mathematics state that in 8th grade, students (and thus also teachers) should be able to "Explain a proof of the Pythagorean theorem and it's converse" (Standard 8.G.6). When discussing the Pythagorean theorem with preservice teachers at a large Midwestern University, the text used and course materials do not differentiate between the theorem and its converse. As such, many preservice teachers in this program are unfamiliar with what the "converse" to the Pythagorean Theorem actually is. As research on learning in online environments suggest that students are more successful in internet searching when they have some understanding of the topic they are investigating (Fidel et al., 1999; Hirsh, 1999), the familiarity with the theorem and the unfamiliarity with the converse makes this standard compelling to investigate..

One way that current research has attempted to classify the ways that students search for information online is by describing the kind of information they are looking for (Bilal, 2000, 2001; Schacter, Chung, & Dorr, 1998) whether it be to find a specific singular answer, or to learn about a topic in a more general sense. Much research on the use of the internet by students suggest that users have a tendency not to question the information they locate or determine whether it comes from a reliable source (Kafai & Bates, 1997; Large & Beheshti, 2000; Lorenzen, 2001).

There have been many different ways of describing the kinds of mathematical knowledge possessed by students. It has been noted recently that the original definitions of procedural and conceptual knowledge described by Hiebert & Lefevre (1986) conflates two separate factors of knowledge, its kind and its quality (Star, 2005). It is argued by Star that knowledge can be either procedural or conceptual in nature, and it can be either rich or superficial in quality. It is with this lens that a more accurate description of the knowledge displayed by learners of mathematics can be obtained.

In this study, two questions are under investigation (1) What form and quality of mathematical connections do preservice teachers make around the Pythagorean Theorem after finding/searching for information online? (2) What are information-seeking strategies that preservice teachers use when finding/searching for information about math tasks online?

Method

Participants in this study focused on improving their own understanding around Common Core State Standard for Mathematics about the proof of the Pythagorean theorem and it's converse, a topic they may be required to teach if they obtained an 8^{th} grade teaching position. They were asked to complete two tasks. During the first task, participants were given open access to the internet to search for and view whatever resources they wanted in order to improve their understanding. In the second task, the participant was directed to 5 specific internet resources and were allowed to explore those resources with the same goal in mind. Before either task began, between tasks, and after both tasks were completed, participants were asked to create a concept map around their understanding of the Pythagorean Theorem.

Each concept map that each participant created was analyzed by first augmenting it with any words they spoke during the map's creation or in describing the map. These augmented maps were divided into smaller chunks for analysis. Each chunk was coded for what kind of mathematical content was included in the chunk, and whether that mathematical content was procedural or conceptual in nature and whether the included content seemed to be of a rich or a superficial quality.

The internet explorations that participants completed were transformed into a database that captured what sites the participant went to, how long they spent on each site, a description of the site's content, a description of the participant's activity while on the site, written and verbal comments made by the participants while on the sites, as well as researcher comments and notes regarding what happened on each site. Participant activity during each task was described by how they appeared to interact and make decisions around sorting through results of internet searches, and how they appeared to interact with the information obtained while examining websites.

Results

Several emergent themes were noted during initial analysis of the maps. First, most all participants did not include the word "converse" in their initial mapping though that word was included in subsequent mappings. Additionally, the majority of the mathematics described appears to be of the superficial quality. Also, participants included a mixture of procedures and concepts in their mappings. One interesting thing to note was how some preservice teachers produced radical shifts in the structuring of later iterations of their concept maps. One student structured her final map in the form of a lesson plan (Figure 1a), describing the order in which she would teach particular content. Another student decided to structure her content around particular themes (Figure 1b).

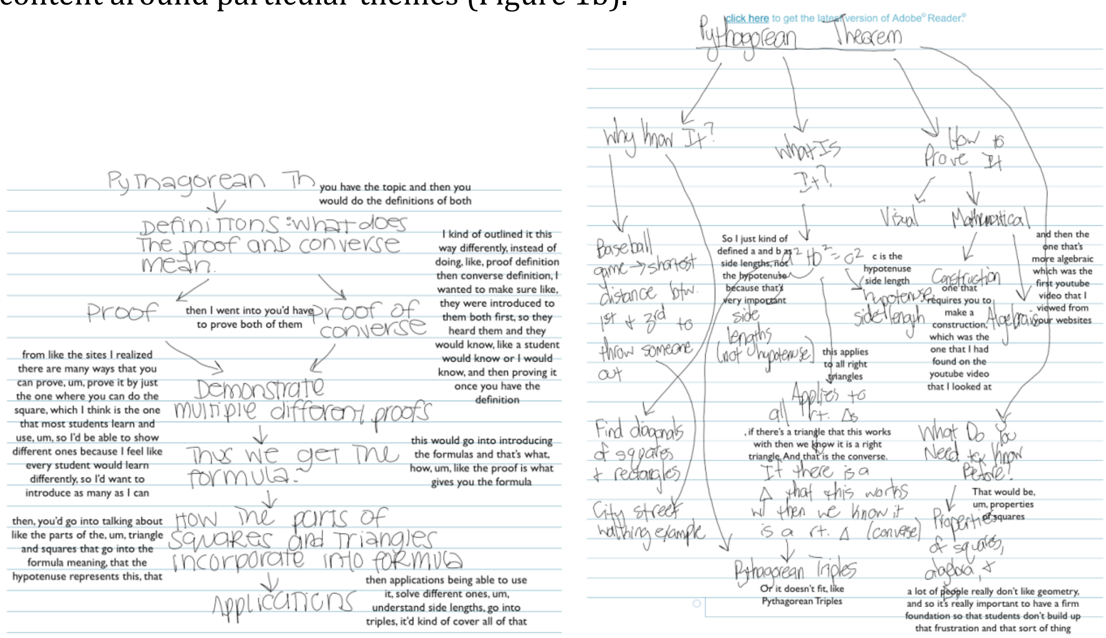


Figure 1(a) and 1(b) – Augmented Concept Maps of Two Participants

Initial results of the internet exploration activities suggest that participants were using the internet to serve different purposes in their explorations. Some were using their explorations as a way of *bookmarking* resources, some were *learning* from websites, and some were *reviewing* information they were already familiar with. When a participant engages in *bookmarking*, they appear to be giving websites a once over, quickly determining if there is information contained in the site that might be of use, and noting that a particular site would be fruitful to explore at a later time for more information. When a participant is *learning* from a website, they appear to be actively reading and even rereading information contained in the site, exploring examples or activities presented in the site, and even working out arguments or examples on their own while engaging with the content of the site. Participants who are *reviewing* content appear to be refamiliarizing themselves with content that they have encountered already. They tend to read portions of or all of a site once before moving onto similar content on other sites. It should be noted that multiple kind of explorations could appear by a participant at different points in their explorations.

There also appear to be two distinct kinds of ways that participants sort through search results, those that are high frequency of sites explored per search, and those that are a low frequency of sites explored through their searching. Those that are high frequency of sites explored tend to have subsequent searches themed towards different kinds of content while those with low frequency of sites explored will often have subsequent searches be around the same content. Additionally, it was not apparent that every student had a way of determining whether information from a website was reliable, though some preservice teachers did make comments noting that particular sites were better ("wolframalpha.com", "any .org site") than others ("wikipedia.com").

Further investigation of these initial results will focus on several questions. First, additional analysis of the content and structure of the concept maps will illuminate how the type and quality of the mathematical information changed (if at all) over the iterations of the concept mapping. Secondly, additional analysis of the internet exploration data will seek to further categorize the kinds of techniques that participants used when investigating mathematical content online. Lastly, any connections that may exist between the kind and quality of the mathematical content included within these mappings and the searching techniques employed by the students will be explored.

During this preliminary research report, attendees will be asked at least three questions to consider while the study and initial results are presented:

- What things might I do to strengthen the argument and to what other literatures does this research speak to?
- What did you find most compelling about the research presented?
- Of the results presented so far, what do you want to know more about?

References

Bilal, D. (2000). Children's Use of the Yahooligans! Web Search Engine: I. Cognitive, Physical, and Affective Behaviors on Fact-Based Search Tasks. *Journal of the American Society for Information Science and Technology*, *51*(7), 646–665.

Bilal, D. (2001). Children's Use of the Yahooligans! Web Search Engine: II. Cognitive and Physical Behaviors on Research Tasks. *Journal of the American Society for Information Science and Technology*, *52*(2), 118–136.

Fidel, R., Davies, R. K., Douglass, M. H., Holder, J. K., Hopkins, C. J., Kushner, E. J., … Toney, C. D. (1999). A visit to the information mall: Web searching behavior of high school students. *Journal of the American Society for Information Science*, *50*(1), 24–37. doi:10.1002/(SICI)1097-4571(1999)50:1<24::AID-ASI5>3.0.CO;2-W

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hirsh, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, *50*(14), 1265–1283. doi:10.1002/(SICI)1097-4571(1999)50:14<1265::AID-ASI2>3.0.CO;2-E

Kafai, Y., & Bates, M. J. (1997). Interet web-searching instruction in the elementary classroom: Building a foundation for information literacy. *School Library Media Quarterly*, 103–111.

Large, A., & Beheshti, J. (2000). The web as a classroom resource: Reactions from the users. *Journal of the American Society for Information Science*, *51*(12), 1069–1080. doi:10.1002/1097-4571(2000)9999:9999<::AID-ASI1017>3.0.CO;2-W

Lorenzen, M. (2001). The land of confusion?: High school students and their use of the World Wide Web for research. *Research Strategies*, *18*(2), 151–163. doi:10.1016/S0734-3310(02)00074-5

Schacter, J., Chung, G. K. W. K., & Dorr, A. (1998). Children's internet searching on complex problems: Performance and process analyses. *Journal of the American Society for Information Science*, *49*(9), 840–849. doi:10.1002/(SICI)1097-4571(199807)49:9<840::AID-ASI9>3.0.CO;2-D

Star, J. R. (2005). Reconceptualizing Procedural Knowledge. *Journal for Research in Mathematics Education*, *36*(5), 404–411.

NAIVE BROUWERIAN VISIONS: A STUDY OF STUDENTS' INTERPRETATIONS OF NON-CONSTRUCTIVE EXISTENCE PROOFS

Stacy Brown
California State Polytechnic University, Pomona

This paper shares findings from a three-phase study exploring students' conceptions of non-constructive existence proofs. Data are used to illustrate students' tendency to apply a naïve Brouwerian lens to non-constructive proofs; that is, a perspective in which learner's proof conceptions are governed by a potentially subconscious anticipation of construction, which enables the learner to construe proofs of existence (be they constructive or non-constructive) as providing actual instances of (or algorithms for producing) mathematical phenomena. Questions concerning researchers proof scheme inferences are raised.

Keywords: Indirect proof, constructive proofs, proof schemes

Researchers have suggested that non-direct proofs, which include proof by contradiction, proof by contraposition, and non-constructive existence proofs, are particularly problematic for students (Tall, 1979; Leron, 1985; Harel & Sowder, 1998; Antonini & Mariotti, 2008). Three different rationales have been provided. First, Tall (1979), who focused specifically on proofs by contradiction, argued that it is the presentation of the proofs that is at issue. Leron also focused specifically on proofs by contradiction and posited that is it the non-constructive nature of such proofs that is problematic for students. Harel and Sowder (1998) provided data from teaching experiments to illustrate students' lack of preference for non-constructive proof and, like Leron, argued that this lack of preference was due to a way of thinking referred to as the constructive proof scheme: a scheme in which "students doubts are removed by the actual construction of objects – as opposed to the mere justification of the existence of objects" (p. 272). Offering an alternative framing, Antonini and Mariotti (2008) argued that students' difficulties with non-direct proofs occur at a meta-theoretical level. In other words, it is difficulties with the acceptance of a logical theory and the resultant move to secondary statements (e.g., from $P \Rightarrow Q$ to $\sim Q \Rightarrow \sim P$ in a proof by contraposition) that is problematic for students. Taken together, these findings indicate that non-direct proofs may be particularly problematic for students.

While providing evidence of students' struggles to accept non-direct proofs, research in this area has not addressed two issues. First, little is known about what might constitute an intellectual need (Harel, 1998) for non-direct proof methods, especially at the undergraduate level. Some evidence in geometric contexts can be found in Antonini and Mariotti (2008) and in numeric contexts with elementary-age children in Maher and Martino (1996). Second, to suggest that students lack a preference for non-constructive proofs is to suggest that students not only comprehend but also reflect on the logical structure of such proofs. Yet, researchers have not documented students' conceptions of non-direct proofs. The purpose of this paper is to make progress on this second issue, so as to identify potential epistemological obstacles (Brousseau, 1997) to non-direct proofs, as part of a program that seeks to address both of the issues listed above. Specifically, the purpose of this paper is to discuss students' conceptions of non-constructive existence proofs and, in particular, to provide evidence indicating that students may struggle to develop an awareness of the logical-structure of non-constructive existence proofs.

A Definition of Non-Constructive Existence Proofs

Theorems in mathematics take many forms. In some cases, theorems assert the existence of a relation, object(s), or other mathematical phenomena. For example, the theorem, "There exists a real-valued solution to the function $f(x) = x^3 - 1$," is an existence statement. Proofs of existence theorems are known as existence proofs. These proofs take two forms: constructive and non-constructive. A *constructive proof* either explicitly provides an instance of a phenomenon or gives an algorithm for generating an actual instance. For example, one can prove the theorem stated above by noting that $(1)^3 - 1 = 0$. In contrast, a *non-constructive proof* demonstrates the logical necessity of a result but fails to produce either a specific instance of a phenomenon or an algorithm for producing a desired result. Non-constructive proofs, therefore, require attention to and use of logical rules of inference. For instance, one can cite the Intermediate Value Theorem, the fact that polynomials are continuous on the reals, and argue that since $f(2) > 0$ and $f(-1) < 0$, then there exists $c \in (-1,2)$ such that $f(c) = 0$. In other words, one can deduce that a solution exists even though the actual value(s) have not been provided and a method for finding the value(s) has not been explicitly stated. With these definitions in mind, it is important to note that it is the former type of proof – constructive – that researchers have suggested students prefer.

Historical Background

Historically, the three forms of non-direct proof were treated differently by the mathematics community. For example, it is well known that Euclid's Elements contains proofs by contradiction and that mathematicians, such as Hardy (1940), felt that this technique was one of mathematicians' "finest weapons." In contrast, in 1890 when Hilbert submitted a paper to the *Mathematische Annalen*, which used a non-constructive existence argument to prove that "if V is a finite dimensional representation of the complex algebraic group $G = \mathrm{SL}n(C)$ then the ring of invariants of G acting on the ring of polynomials $R = S(V)$ is finitely generated," his paper was met with significant objections. Indeed, Paul Gordan, a leader in invariant theory, is reported to have responded, "This is not mathematics. It is theology" (Webb, 1997, p. 1). Writing to Felix Klein, Gordan argued, "Hilbert has scorned to present his thoughts following formal rules; he thinks it suffices that no one contradicts his proof, then everything will be in order … he thinks that the importance and correctness of propositions suffice … but of a comprehensive work for the *Annalen* this is insufficient" (Rowe, 1986, cited in Webb, 1997, p. 1) Hilbert's proof had established the logical necessity of a finite basis without constructing such a basis.

Later, in 1909, L. E. J. Brouwer became famous for results in the field of point set topology and most specifically for a result known as the Brouwer Fixed Point Theorem. This result, like Hilbert's proof of a finite basis, used a non-constructive existence proof to established that "for any continuous function f with certain properties mapping a compact convex set into itself, there is a point x_0 such that $f(x_0) = x_0$." Later, however, Brouwer became an adamant intuitionist and, consequently, came to view his own proof of the Brouwer Fixed Point Theorem as an invalid proof of existence (within intuitionistic logic) and to seek a constructive proof. Indeed, Brouwer progressively adopted a philosophy of mathematics in which, "objects of mathematics are mental creations, and hence that they can be said to exist if and only if those creations have actually been carried out" (Engel, 2007, p. 1). Since logical necessity arguments do not construct the mathematical objects whose existence they seek to establish, Brouwer rejected such proofs as proofs of existence. Specifically, arguments of logical necessity, such a Hilbert's, were viewed as proving the statement ~P is false, where P is the statement of existence, but not as proving P since a proof of existence requires something more; namely, a construction. (It is due to this and other points of logic that the Intuitionists were viewed as rejecting the law of the excluded middle.) Thus, though much more can be said, it is clear from the historical events reported that consideration of non-constructive existence proofs created significant philosophical issues for the mathematics community and, in some cases, ruptures.

Theoretical Considerations

To *conceive* means "to form or develop" whereas *perceive* means "to attain awareness or understanding of." One may conceive of a proof in two ways: (1) the individual develops a proof through actions we would not classify as repetitions or reiterations of prior activities, or (2) the individual is presented with a proof and develops an image of the proof in their mind. In the later case, the individuals' ways of understanding mathematical constructs and ways of thinking about mathematics will afford and constrain the meanings available to the learner. Ways of understanding can be thought of as "the particular meaning/interpretation a person gives to a concept, relationships between concepts, assertions, or problems" (Harel & Sowder, 2005, p. 30). Ways of thinking "refers to what governs one's ways of understanding, and thus expresses reasoning that is not specific to one particular situation but to a multitude of situations" and involves a person's "beliefs, problem-solving approaches and proof schemes" (p. 31).

Proof schemes are "what constitutes ascertaining and persuading for that person" (Harel & Sowder, 2005, p. 33). More often than not, researchers infer learners' proof schemes by evaluating the proofs the students produce and assuming that the students' interpretations of the produced proofs parallel the researchers' interpretations. Some have implied that this methodological approach may be problematic. For instance, researchers interested in understanding students' use of empirical arguments have begun to question what can be ascertained when students produce such arguments. Indeed, Weber (2009) and Vinner (1997), have argued that students may produce empirical arguments if they are unable to produce a proof but wish to produce something. Others (Brown, *submitted*) have questioned if students' exploratory work has been misclassified as an attempted proof.

An alternative to examining students' proof productions is to explore students' conceptions of presented proofs. In particular, by asking students to explain a given argument researchers may be able to infer the criteria students bring to bear on proofs and to separate out instances in which students' lack the necessary ways of understanding from those instances in which the learner's ways of thinking inhibit or promote students' proof conceptions. This approach reaches back to the Piagetian roots of the term *scheme* and draws on Thompson's definition: "A scheme is an organization of actions that has three characteristics: an internal state that is necessary for the activation of actions composing it, the actions themselves, and an imagistic anticipation of the result of acting" (Thompson, 1994, p. 182). Here we see that in addition to organizing action, schemes play an anticipatory role. Thus, from this perspective, one can argue that a learner's proof schemes may not only influence the proofs produced but may also play a role in the outcomes anticipated as students seek to interpret or produce a proof.

The outcomes anticipated by a learner may be either consciously or subconsciously attended to by a learner. In the former case, the learner can explicitly state the criteria they bring to bear on a proof. In the latter case, the learner may be unaware of these criteria for the criteria are part of the learner's tacit knowledge – ways of knowing that "act and influence the reasoning process without the individual being aware of their origin and their effect" (Fischbein, 2001, p. 313). When applied to contexts of interpretation, such anticipations could be thought of as anticipatory intuitions rather than intuitive affirmations (Fischbein, 1983), for it is a feeling of *what is to be* rather than a feeling of "it must be so" that is responded to by the leaner. Thus, how a learner *conceives* of a proof can be seen as indicative of learners' proof schemes, which are viewed as governing the learner's anticipatory behaviors at either a conscious or subconscious level. These schemes influence not only learners' production but also their reproduction of mathematical proofs and, consequently, are indicated by learners' modifications during acts of reproduction – modifications the learner may be aware of at either a conscious or subconscious level.

The Study

This paper reports findings from a three-stage study of students' perceptions of and preferences for (or lack of preference for) *non-direct mathematical proofs*; that is, non-constructive existence proofs, proofs by contraposition, and proof by contradiction. The first stage of the study involved administering a written survey to mathematics majors ($n = 21$) who were either completing an introduction to proof course or enrolled in an advanced topic course (analysis, topology, etc.). Results from the survey were reported in (Brown, 2011). Briefly, the comparison tasks asked students to compare a direct proof to one of the three types of non-direct proof and to indicate: (a) their level of confidence; (b) the most convincing proof; and, (c) the best proof. Despite prior observations indicating that students' prefer direct proofs (Leron, 1985; Harel & Sowder, 1998), data from the first stage of the study did not indicate a lack of preference for non-direct proofs. For example, Task 3 asked students to compare a direct proof to a proof by contraposition. Of the 21 responses, 9 ranked the direct proof as most convincing and 12 chose the proof by contraposition. With regard to the comparison of a non-constructive existence proof with a constructive proof (Task 4), 9 students found the constructive proof "most convincing," while 13 chose the non-constructive proof (See Appendix A for Tasks 3 and 4).

The second stage of the study involved 6 clinical interviews with mathematics majors, who were selected to represent a variety of survey responses. During the interviews, students discussed their survey responses and responded to comprehension questions. The goal of the clinical interviews was to create case studies linking students' interpretations of the arguments to their proof preferences. The third stage of the study involved a classroom intervention ($n = 29$), which included activities intended to foster growth in students' perceptions and production of non-constructive existence proofs. The intervention employed "conceptual awareness pillars" (Stylianides & Stylianides, 2009); i.e., questions aimed at developing students' awareness of the logical features of non-constructive existence proofs. Students were given an end-of-term assessment, which included a reproduction item; that is, students were asked to prove an existence statement that had been discussed in class and proven using a non-constructive existence proof. In this paper, findings from the second and third stages of the study are reported.

Findings

Analyses of the clinical interview transcripts indicated that 4 of the 6 students struggled to comprehend the logical structure of Argument B in Task 4 (see Appendix A), which I will refer to as Argument 4B. Due to space limitations, however, we will focus on students' responses to the interview question, "Does the argument provide a specific case for which the theorem is true?" The two students (all names are pseudonyms), who successfully responded to the comprehension questions, responded to this question by noting that the argument did not provide a specific instance of the theorem. For example, Nora argued, "Here [Argument B], we found an *a* and *b* and we said, well, there's this one case that might be true and then there's this other case that might, that … that one of these two things is true. […] So, either, either this pair of numbers is applicable or this pair of numbers is applicable." In contrast, 3 students argued that either one or two specific instances of the theorem were provided in the proof. For example, Jake, who selected Argument 4B as most convincing, argued that a specific instance had been provided in the proof (see Figure 1). In other words, these students' interpretations of the argument were constructive. The remaining student initially argued that a specific instance was provide and then after a period of time spent considering the question, “Does the argument provide a specific case for which the theorem is true,” self-corrected and provided a description similar to Nora’s.

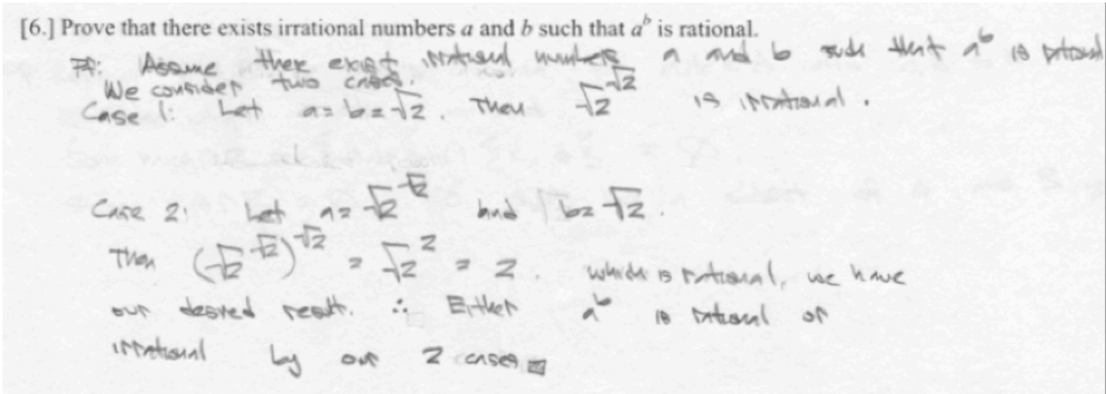
Int: Umm, so at the end of Argument B, do we have an ***a*** and a ***b*** that are irrational, such that ***a*** to the ***b*** is rational? Do we have a set of numbers?
Jake: Well, we have one example. We say ***a*** is the square root of two and ***b*** is the square root of two.
Int: Okay.
Jake: And, this proof shows ***a*** to the ***b*** is two, which is rational.

Figure 1. Jake’s Transcript

Discussion of findings from the third stage of the study will be restricted (due to space limitations) to students’ in-class responses to the non-constructive existence proof of Theorem 4 (see Appendix A) and to a reproduction task involving Theorem 4 on the end-of-term assessment. During the classroom intervention approximately 45% of the students (13 out of 29) indicated with gestures (raised hands) that Argument 4B produced at least one specific instance of the theorem. In other words, nearly half of the students viewed the proof constructively. Inquiries into the students’ interpretations highlighted students’: (1) assumption that the number $\sqrt{2}^{\sqrt{2}}$ is irrational, and (2) difficulties with conditional statements. Use of the assumption “$\sqrt{2}^{\sqrt{2}}$ is irrational” led to a whole-class discussion about the lack of means (previous theorems or proof techniques) with which to classify this number as either rational or irrational.[1]

Four codes were used to categorize student responses to the end-of-term assessment item, which was a proof reproduction task involving Theorem 4: (a) valid, non-constructive existence proof (44.8%); (c) invalid, direct proof (31%); (e) invalid, proof by contradiction (10%); and (f) other (14%). Proofs of type (a) tended to be reproductions of Argument 4B. The remainder of this discussion will focus on proofs of type (c). Responses were coded as an *invalid, direct proof* if the student choose specific values for a and b and attempted to shown that a^b is rational. Comparison of the students’ attempted direct proofs to the previously shown non-constructive proof indicated that the attempted direct proofs were markedly similar in two distinct ways; they were either a *concatenation* or a *redaction* of the non-constructive proof. An attempted direct proof was classified as a *concatenation* if the student: (a) did not use the law of the excluded middle, and (b) replaced the conditional statements in Argument 4B with a set of dependent, declarative sentences. This type of attempted direct proof (Figure 2) was called a concatenation because the original conditional statements were revised and “strung together” to create a single case. Approximately 44% of the attempted direct proofs were concatenations.

[1] The fact that number $\sqrt{2}^{\sqrt{2}}$ is irrational is a consequence of the Gelfond-Schneider Theorem (GST), which is otherwise known as *Hilbert’s Seventh Problem*. The proof of the GST was viewed as beyond the scope of the introductory class and consequently, the irrationality of $\sqrt{2}^{\sqrt{2}}$ was viewed as unproven.


[6.] Prove that there exists irrational numbers a and b such that a^b is rational.

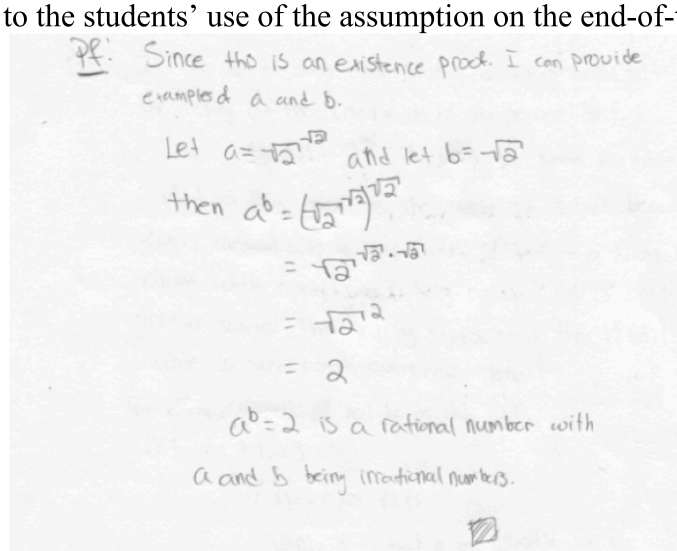
Pf: Assume there exist irrational numbers a and b such that a^b is rational
We consider two cases.
Case 1: Let $a=b=\sqrt{2}$. Then $\sqrt{2}^{\sqrt{2}}$ is irrational.

Case 2: Let $a=\sqrt{2}^{\sqrt{2}}$ and $b=\sqrt{2}$.
Then $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{2} = 2$. which is rational, we have our desired result. ∴ Either a^b is rational or irrational by our 2 cases ∎

Figure 2. Concatenation Example.

An argument was classified as a *redaction* if the attempted direct proof could be viewed as an edited-version of the non-constructive proof, where a significant portion of the argument was removed. Specifically, these attempts involved the omission of first 4 lines of Argument 4B and the replacement of the conditional clause, "if $\sqrt{2}^{\sqrt{2}}$is irrational" with declarative statements about the irrationality of either $\sqrt{3}^{\sqrt{2}}$ or $\sqrt{2}^{\sqrt{2}}$. Redactions comprised 55% of the attempted direct proofs.

There are two plausible explanations for students' production of redactions (see Figure 3) and concatenations of the non-constructive proof. The first explanation is that, despite the explicit in-class discussions, students persisted in their use of an unproven assertion ($\sqrt{2}^{\sqrt{2}}$ is irrational) because the assertion was intuitively acceptable. Though plausible, evidence of this explanation is limited to the students' use of the assumption on the end-of-term assessment.


Pf: Since this is an existence proof. I can provide examples of a and b.

Let $a=\sqrt{2}^{\sqrt{2}}$ and let $b=\sqrt{2}$

then $a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$

$= \sqrt{2}^{\sqrt{2}\cdot\sqrt{2}}$

$= \sqrt{2}^{2}$

$= 2$

$a^b = 2$ is a rational number with a and b being irrational numbers.

∎

Figure 3. Redaction Example

A second explanation is that the students interpreted Argument 4B using a *Naïve-Brouwerian lens*: a perspective in which learner's proof conceptions are governed by a potentially subconscious anticipation of construction, which enables the learner to construe proofs of existence (be they constructive or non-constructive) as providing actual instances of (or algorithms for producing) mathematical phenomena.[2] Evidence of this explanation can be found

[2] The term "naïve" in the phrase *Naïve-Brouwerian lens* is used intentionally. Brouwer's position on non-constructive arguments was rooted in a deep understanding of non-constructive proofs and their logical underpinnings. In the students' case, an anticipation of construction rather than an awareness of a lack of construction guides their interpretations. Hence, they are naïve – they lack awareness.

both in the clinical interviews and the in-class field notes. For example, in one interview, Maria proposed that the argument created a new value for *a*, an irrational number ($\sqrt{2}^{\sqrt{2}}$), “so that you get a rational when you put *a* to the *b*.” In other words, she interpreted the statements of the proof as “building” specific numbers, which could then be employed in the service of a constructive proof. These data suggest that it was an anticipation of construction, rather than the preservation of an assumption, that governed students’ reproductions of the non-constructive proof and resulted in the redacted and concatenated versions described above.

Discussion and Implications

Tall (1991) noted that many researchers (Ellerton, 1985; Biggs & Collins, 1982) have suggested that in advanced mathematics, the developmental stages identified by Piaget may be viewed as a learning cycle occurring at each individual stage and producing, though each cycle, “a higher level of abstraction” (p. 8). If this is the case, then there may be a concrete-operational phase within the hypothetical-deductive stage. Thought of in relation to proof schemes, learners’ initial attempts to conceive of non-constructive proofs may be hindered by leaner’s constructive anticipations; that is, governed by a *Naïve-Brouwerian lens*. However, more research is need within other proof situations, which avoid intuitively acceptable assumptions, if researchers are to determine if such a perspective is, in fact, dominant among novices.

The implications of this study are two fold. First, findings from the survey and clinical interviews indicate that researchers may need to document not only students’ proof productions but also students’ interpretations of their own and others’ proof productions, if we are to better understand students’ proof schemes. Such work would, among other methodological changes, involve including comprehension questions such as those proposed by Mejia-Ramos, Fuller, Weber, Rhoads, & Samkoff (2012). Indeed, the clinical interviews indicate that students’ selection of Argument 4B was impacted by a lack of comprehension of the argument. Second, findings from the intervention indicate that non-constructive existence proofs may pose particular problems for novices as they attempt to interpret these proofs. Consequently, instructional interventions intended to advance novices’ production and comprehension of such proofs should afford opportunities for students to become aware of and explicitly discuss the metatheorical basis for such proofs.

References

Antonini, S., & Mariotti, M. A. (2008). Indirect Proof: What is specific to this way of proving? *ZDM – The International Journal on Mathematics Education*, *40*, 401 – 412.

Biggs, J. & Collins, K. (1982). *Evaluating the Quality of Learning: The SoloTaxonomy*. New York, NY: Academic Press.

Brousseau, G. (1997). *Theory of Didactical Situations in Mathematics*. Dordrecht, Netherlands: Kluwer

Brown, S. *On Skepticism and It’s Role in the Development of Proof in the Classroom*. Submitted

Brown, S. (2011, February) An Investigation of Students’ Proof Preferences: The Case of Indirect Proofs. In S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman (Eds.), *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education,* Vol. 4, p. 31 – 36, Portland, Oregon.

Ellerton, N. F. (1985). *The Development of Abstract Reasoning – Results from a Large Scale Study in Australia and New Zealand*.

Engel, P. (2003). *Dummet, Achilles and the tortoise*. [To appear in The Philosophy of Michael Dummett, The Library of Living Philosophers] Electronic retrieval December 15, 2013.

Fischbein, E. (1983). Intuition and proof. *For the Learning of Mathematics*, 3(2), 9 – 18.
Fischbein, E. (2001). Tacit models and infinity. *Educational Studies in Mathematics*, 48(2/3), 309 – 329.
Hardy, G. H. (2004) [1940]. *A Mathematician's Apology*. Cambridge: University Press.
Harel, G. (1998). Two Dual Assertions: The First on Learning and the Second on Teaching (or Vice Versa). *American Mathematical Monthly*.105(6), 497-507.
Harel, G., & Sowder, L. (1998). Students' Proof Schemes. In E. Dubinsky, A. Schoenfeld, and J. Kaput (Eds.) *Research on Collegiate Mathematics Education*. 234-283. USA: American Mathematical Society.
Leron, U. (1985). A direct approach to indirect proofs. *Educational Studies in Mathematics*, *16(3),* 321 – 325.
Maher, C. A. & Martino, A. M. (1996). The development of the idea of mathematical proof: A 5-year case study. *Journal for Research in Mathematics Education, 27(2),* 194-214.
Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics, 79*(1), 3-18.
Stylianides, A., & Stylianides, G. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40(3),* 314-352.
Tall, D. (1991). The psychology of advanced mathematical thinking. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 3 – 21). Dordrecht, Netherlands: Kluwer.
Thompson, P. (1994). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 181 – 234). Albany, NY: SUNY Press.
Webb, J. (1997). Hilbert's formalism and arithmetization of mathematics. *Synthese,* A *Symposium on David Hilbert*, 110(1), 1 – 14.
Weber, K. (2009). Mathematics Majors' Evaluation of Mathematical Arguments and Their Conception of Proof, *Electronic Proceedings for the 12th Annual Conference on Research in Undergraduate Mathematics Education*: Raleigh, NC.
Vinner, S. (1997). The pseudo-conceptual and pseudo-analytic thought processes in mathematical

Appendix A.

Theorem 3: Suppose a set A has the property, for any subset B, $A \subseteq B$, then $A = \emptyset$.

Argument A:	**Argument B:**
Let B be an arbitrary subset of A, then $A \subseteq B$. Let $B^* = A \cap \overline{B}$. B^* is a subset of A. It follows that $A \subseteq B^*$. Since $A \subseteq B^*$ and $A \subseteq B$, $A \subseteq B \cap B^*$. However, $B \cap B^* = \emptyset$. Hence, $A = \emptyset$.	Suppose $A \neq \emptyset$. Since $A \neq \emptyset$ there is an element x, such that $x \in A$. Now, let $B = \emptyset$. Thus, B is a subset of A such that $x \notin B$. It follows that $A \not\subseteq B$, as desired.
1. I am confident about my understanding of Argument A. (Please mark one) ☐ Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree	**2. I am confident about my understanding of Argument B. (Please mark one)** ☐ Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree

3. Which argument, in your opinion, is the most convincing? ☐ Argument A ☐ Argument B

Please explain your selection. (If you need additional space please use the back of this page.)

4. Which argument, in your opinion, is the best proof? ☐ Argument A ☐ Argument B

Please explain your selection. (If you need additional space please use the back of this page.)

Theorem 4: There exist irrational numbers a and b such that a^b is rational.

Argument A

<u>*Definition 1*</u>. $\log_z x$ is the number y, such that $x = z^y$.

<u>*Lemma 1.0*</u>: The $\log_2 9$ is irrational.
Proof: Suppose $\log_2 9$ is rational.
Since $\log_2 9$ is rational, there exists then integers m and n such that $\log_2 9 = \frac{m}{n}$.

By *Definition 1.*, it follows that $2^{\frac{m}{n}} = 9 \Rightarrow 2^m = 9^n$
However, 9^n is odd, for all integers n, and 2^m is even, for all integers m, which is a contradiction. Thus, $\log_2 9$ is irrational.

<u>*Theorem 4*</u>: There exist irrational numbers a and b such that a^b is rational.
Proof: Let $a = \sqrt{2}$. It is well known that the $\sqrt{2}$ is irrational. Let $b = \log_2 9$, by *Lemma 1.0*, b is irrational.

$$a^b = \sqrt{2}^{\log_2 9} = (2^{\frac{1}{2}})^{\log_2 9} = (2^{\log_2 9})^{\frac{1}{2}} = (9)^{\frac{1}{2}} = 3$$

Thus, a^b is rational.

Argument B

Let $a = \sqrt{2}$ and $b = \sqrt{2}$.
It is well known that $\sqrt{2}$ is irrational.

$$a^b = \sqrt{2}^{\sqrt{2}}$$

If $\sqrt{2}^{\sqrt{2}}$ is rational, then the theorem is true.

If $\sqrt{2}^{\sqrt{2}}$ is irrational, let $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$, then

$$a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = (\sqrt{2})^{\sqrt{2}\cdot\sqrt{2}} = (\sqrt{2})^2 = 2.$$

Thus, a^b is rational.

1. I am confident about my understanding of Argument A. (Please mark one)

☐ Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree

2. I am confident about my understanding of Argument B. (Please mark one)

☐ Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree

3. Which argument, in your opinion, is the most convincing? ☐ Argument A ☐ Argument B
Please explain your selection. (If you need additional space please use the back of this page.)

4. Which argument, in your opinion, is the best proof? ☐ Argument A ☐ Argument B

UNDERGRADUATE STUDENTS' USE OF INTUITIVE, INFORMAL, AND FORMAL REASONING TO DECIDE ON THE TRUTH VALUE OF A MATHEMATICAL STATEMENT

Kelly M. Bubp
Ohio University

Although deciding on the truth value of mathematical statements is an important part of the proving process, students are rarely engaged in making such decisions. Thus, little is known about the ways in which students use intuitive, informal, and formal reasoning to evaluate conjectures. In this study, task-based interviews were conducted with undergraduate students in which they were asked to determine the truth value of five mathematical statements on functions and relations. Students' reasoning on these tasks will be classified as intuitive, informal, or formal, and then further categorized according to the findings of current research, with new categories added as needed. This study should contribute to our understanding of the ways in which students reason when dealing with uncertainty in the proving process. Additionally, this study may suggest ways in which educators can assist students in navigating the often difficult process of proving and refuting mathematical statements.

Key words: Reasoning and proof, Evaluating conjectures, Intuition, Task-based interviews

The process of proving mathematical statements often starts with formulating or evaluating mathematical conjectures—activities that involve uncertainty. Unfortunately, students rarely are engaged in uncertain aspects of the proving process, including determining the truth value of mathematical statements (Alibert & Thomas, 1991; de Villiers, 2010; Durand-Guerrier, Boero, Douek, Epp, & Tanguay, 2012). Due to these limited opportunities, little is known about how students evaluate conjectures and what types of reasoning they use to do so. In particular, there is a lack of research on students' ways of deciding on the truth value of general mathematical statements involving general mathematical objects in the context of proof-based mathematics.

Determining the truth value of mathematical statements is an important component of the proving process. When dealing with uncertainty, mathematicians often try to decide on a statement's truth value with some degree of confidence before investing time in a proof or refutation attempt (de Villiers, 2010; Inglis, Mejia-Ramos, & Simpson, 2007). Study of this decision process is essential for determining the ways in which intuitive, informal, and formal reasoning can be used to make successful decisions about the truth or falsity of mathematical statements. Furthermore, study of successful students engaging in this process, rather than mathematicians, is more likely to yield results of pedagogical value and "suggest learning trajectories that might be applicable for many other students as well" (Weber, 2009, p. 201). Thus, this study will explore the following question: In what ways and to what extent do undergraduate students use intuitive, informal, and formal reasoning to decide on the truth value of a mathematical statement?

Literature Review

Intuitive, informal, and formal reasoning are three types of reasoning which can assist with the process of deciding on the truth value of a mathematical statement. *Intuition* constructs an automatic mental representation of a task, taking into consideration task cues, prior knowledge, and experience, and operates independently of working memory (Evans, 2010, 2012; Fischbein, 1987; Glockner & Witteman, 2010; Wilder, 1967). Informal and formal reasoning are deliberate processes of reasoning that can be explained, decomposed into their constituent parts, and require the use of working memory (Evans, 2008, 2012; Fischbein, 1987). *Informal* reasoning includes a variety of reasoning strategies such as visuo-spatial, example-based, graphical, diagrammatic, physical, kinaesthetic, analogical, inductive, and pattern-based. *Formal reasoning* is reasoning from definitions, axioms, assumptions, and theorems based solely on logic and deduction that conforms to specified rules regarding language, symbols, and frameworks for argumentation. Much of the research on the types of reasoning engaged in while deciding on the truth value of a mathematical statement comes from the study of mathematicians' reasoning rather than students' reasoning.

Intuitive Reasoning

Intuition is especially important for deciding on the truth value of a mathematical statement because it can suggest what is plausible in the absence of a proof (Burton, 2004; Davis & Hersh, 1981; Fischbein, 1994) and "provides a justification for, but is prior to, the search for convincing argument and, ultimately, proof" (Burton, 1999, p. 32). In the limited research on intuition in mathematics education, researchers have found a variety of types of intuitive reasoning used by students and mathematicians to evaluate mathematical conjectures. Inglis et al. (2007) found that mathematicians' intuitive support for the truth or falsity of a mathematical statement was based on either suspected properties about mathematical objects or known relationships between mathematical concepts. On the other hand, students' intuitive decisions on the truth value of mathematical statements have been found to be based on (a) cues in the statement (Buchbinder & Zaslavsky, 2007; Leron & Hazzan, 2006); (b) definitions and mental images (Bubp, 2012); or (c) expected relationships between mathematical objects with certain properties (Bubp, in press).

Informal Reasoning

Mathematicians use a variety of informal reasoning strategies to convince themselves or to reduce their uncertainty of the truth value of a mathematical statement, including: drawing geometric figures or diagrams (de Villiers, 2010), examining special or limiting cases (Alcock & Inglis, 2008; de Villiers, 2010), reasoning by analogy (Alcock & Inglis, 2008; de Villiers, 2010), exploring patterns (Alcock & Inglis, 2008; de Villiers, 2010), studying specific or generic examples (Alcock & Inglis, 2008; Inglis et al., 2007), searching for counterexamples (Inglis et al., 2007), or engaging in informal plausibility argumentation about properties of relevant mathematical objects (Alcock & Inglis, 2008; Inglis et al., 2007). Although much less is known about how students reason informally about the truth value of a mathematical conjecture, evidence has been found that they (a) study specific examples (Buchbinder & Zaslavsky, 2007; Connor, Moss, & Grover, 2007; Durand-Guerrier et al., 2012; Weber & Mejia-Ramos, 2009); (b) search for counterexamples (Durand-Guerrier et al., 2012); (c) draw diagrams, especially Venn diagrams (Weber, Brophy, & Lin, 2008); and (d) construct informal arguments about mathematical properties of examples (Weber & Mejia-Ramos, 2009).

Formal Reasoning

Mathematicians and students both employ various formal reasoning strategies to decide on the truth value of a mathematical statement. Inglis et al. (2007) found that mathematicians reasoned formally from definitions, algebra, and counterexamples to determine the truth value of a mathematical conjecture. Weber (2009) provides an account of a successful undergraduate student who used only formal reasoning to evaluate conjectures. This student would reformulate a conjecture by using logically equivalent statements or alternate definitions, determine logical inferences that could be made from the assumptions, or attempt a proof. Additionally, students may perform algebraic or symbolic manipulations (Buchbinder & Zaslavsky, 2007) or consider possibly relevant theorems, rules, or definitions when determining the truth value of a mathematical statement (Buchbinder & Zaslavsky, 2007; Durand-Guerrier et al., 2012).

Method of Inquiry

Participants

Purposeful sampling was used to select participants from the main campus of a public university in Ohio. All selected participants met the criteria of (a) being an undergraduate student enrolled in selected mathematics courses at the university, and (b) having passed at least one proof-based mathematics course with a B or better as an undergraduate student.

Procedures

I conducted two task-based interviews with each participant which were audio-recorded and transcribed. Participants were asked to think aloud during completion of the tasks and to clarify or expand on their thinking as necessary. The tasks were provided one at a time on separate sheets of paper. Participants were not (a) given any instructions other than to think aloud during the process, (b) provided with any information other than a list of definitions of terms in the tasks, or (c) given the use of any materials other than a LiveScribe Pen and paper. This Pen records synchronously audio and writing. After each task, I asked follow-up questions regarding the participants' work on the task and general questions on their approach to uncertainty in the proving process.

Tasks

Participants completed five tasks in which they were asked to determine the truth value of a given mathematical statement and prove or disprove the statement accordingly. The tasks cover basic information about functions and relations with which students who have taken at least one proof-based mathematics course should be familiar, thus they have been chosen to be accessible to the participants. In line with Alcock and Weber (2010), each of the tasks refers to general objects and their properties and should be amenable to intuitive, informal, or formal reasoning strategies. Finally, the tasks provide opportunities to construct both proofs and counterexamples. The following are two of the tasks used in the study:

Monotonicity task: Prove or disprove: If $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ decreasing on an interval *I*, then the composite function $f\circ g$ is increasing on *I*.

Composite function task. Let $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ be functions. Prove or disprove: If the composite function $f\circ g$ is one-to-one, then g is one-to-one.

Analysis

Reasoning used during the process of deciding on the truth value of the given mathematical statements will be classified as either intuitive, informal, or formal. I will classify reasoning as intuitive if the student (a) stated that it was a(n) intuition, instinct, gut feeling, or first thought; (b) used similarity to make an assessment of the task; or (c) was unable to justify the reasoning. Reasoning will be classified as informal if it involves deliberate and justifiable use of: visuo-spatial, example-based, graphical, diagrammatic, physical, kinaesthetic, analogical, inductive, or pattern-based reasoning. Reasoning will be classified as formal if it involves deliberate and justifiable logical and deductive reasoning based on axioms, definitions, theorems, given assumptions, and standard proof frameworks. After the initial classification into intuitive, informal, and formal, the reasoning will be further categorized according to the types discussed in the literature review, with new categories added as necessary. Each instance of reasoning will be classified separately in order to capture the use of multiple reasoning types throughout the decision process.

Preliminary Results

Analysis is on-going, but preliminary findings indicate that although the participants in this study had difficulty deciding whether an open statement was true or false, they used a combination of intuitive, informal, and formal reasoning to assist them in making this decision. Participants' intuitive reasoning includes using vague images of functions with particular properties or considering what makes sense based on prior mathematical experience. Their informal reasoning includes exploring specific examples and sketching graphs. Finally, participants' formal reasoning involves reasoning from definitions or making logical implications to see if these lead to information that could support a decision.

Questions for the Audience

Are there other frameworks for analyzing students' reasoning during the process of deciding on the truth value of a mathematical statement that may be more informative than the one I am using? How might I improve my current framework? What educational implications do you see from this work?

References

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69,* 111–129.

Alcock, L., & Weber, K. (2010). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. In F. Hitt, D. A. Holten, & P. Thompson (Eds.), *Research in collegiate mathematics education. VII* (pp. 93–114). Providence, RI: American Mathematical Society.

Alibert, D., & Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 215–230). Dordrecht, Netherlands: Kluwer Academic.

Bubp, K. (2012). To prove or disprove: In what ways do undergraduate students use intuition to decide? In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Kalamazoo, MI: Western Michigan University.

Bubp, K. (in press). Systematic intuitive errors on a prove-or-disprove monotonicity task. To appear in the *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education*. Denver, CO.

Buchbinder, O. & Zaslavsky, O. (2007). How to decide? Students' ways of determining the validity of mathematical statements. In D. Pitta-Pantazi, & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 561–570). Larnaca, Cyprus.

Burton, L. (1999). Why is intuition so important to mathematicians but missing from mathematics education? *For the Learning of Mathematics, 19*(3)*,* 27–32.

Burton, L. (2004). *Mathematicians as enquirers: Learning about learning mathematics*. Boston, MA: Kluwer.

Connor, J., Moss, L., & Grover, B. (2007). Student evaluation of mathematical statements using dynamic geometry software. *International Journal of Mathematical Education in Science and Technology, 38,* 55–63.

Davis, P. J. & Hersh, R. (1981). *The mathematical experience*. Boston, MA: Birkhauser.

de Villiers, M. (2010). Experimentation and proof in mathematics. In G. Hanna, H. N. Jahnke & H. Pulte (Eds.) *Explanation and proof in mathematics: Philosophical and educational perspectives* (pp. 205–221). New York, NY: Springer.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S., & Tanguay, D. (2012). Argumentation and proof in the mathematics classroom. In G. Hanna & M. de Villiers (Eds) *Proof and proving in mathematics education: The 19th ICMI study* (pp. 349–367). Dordrecht, Netherlands: Springer.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59,* 255–278.

Evans, J. St. B. T. (2010). Intuition and reasoning: A dual-process perspective. *Psychological Inquiry, 21,* 313–326.

Evans, J. St. B. T. (2012). Spot the difference: Distinguishing between two kinds of processing. *Mind & Society, 11,* 121–131.

Fischbein, E. (1987). *Intuition in science and mathematics: An educational approach*. Dordrecht, Netherlands: Kluwer.

Fischbein, E. (1994). The interaction between the formal, the algorithmic, and the intuitive components in a mathematical activity. In R. Biehler, R. W. Scholz, R. Straser, & B. Winkelmann (Eds.) *Didactics of mathematics as a scientific discipline* (pp. 231–245). Dordrecht, Netherlands: Kluwer.

Glockner, A. & Witteman, C. (2010). Beyond dual-process models: A categorization of processes underlying intuitive judgment and decision making. *Thinking and Reasoning, 16,* 1–25.

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modelling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66,* 3–21.

Leron, U. & Hazzan, O. (2006). The rationality debate: Application of cognitive psychology to mathematics education. *Educational Studies in Mathematics, 62,* 105–126.

Weber, K. (2009). How syntactic reasoners can develop understanding, evaluate conjectures, and generate counterexamples in advanced mathematics. *Journal of Mathematical Behavior*, *28*, 200–208.

Weber, K., Brophy, A., & Lin, K. (2008). Learning advanced mathematical concepts by reading text. In *Proceedings of the 11th conference for research in undergraduate mathematics education*. Retrieved from http://sigmaa.maa.org/rume/crume2008/Proceedings/Weber%20LONG.pdf

Weber, K., & Mejia-Ramos, J. P. (2009). An alternative framework to evaluate proof productions: A reply to Alcock and Inglis. *Journal of Mathematical Behavior*, *28*, 212–216.

Wilder, R. L. (1967). The role of intuition. *Science*, *156*, 605–610.

SUPPORTING STUDENTS TO CONSTRUCT PROOFS: AN ARGUMENT ASSESSMENT TOOL

Justin Boyle
University of New Mexico

Martha Byrne
University of New Mexico
Earlham College

Engaging students in the construction of proofs often does not include conversations about what does and does not count as proof, to the detriment of the students. The critiques of student-generated arguments should be communicated in a language common to instructor and student; such a language can be developed via an assessment tool that is accessible to both parties. This paper describes the development of an argument assessment tool that will be useful for instructors and researchers both to assess students' and participants' ability to construct proofs and to communicate those assessments. The tool is introduced and two assessed student arguments are shared to illustrate the tool's application. Future work with the argument assessment tool will include its use in a classroom as an instructional tool for establishing a common language for instructor and students and providing the foundation for discussions about proof production.

Key words: Mathematical Proof; Argument Assessment; Constructivism

Introduction

There exists extensive research on undergraduate mathematics students' limited understanding with respect to constructing and evaluating proofs. The arguments for why undergraduate students struggle vary from not knowing how to start a proof (Moore, 1994), to a contrived understanding of mathematical definitions (Ewards & Ward, 2004; Moore, 1994), an inability to move flexibly between inductive and deductive thinking (Ball, Hoyles, Jahnke & Movshovitz-Hadar, 2002), and a lack of understanding of the difference between inductive and deductive arguments (Morris, 2002). A proposed solution to this problem is the transformation of instruction from traditional lecture-based courses to constructivist classrooms where students work intentionally to create their own knowledge. Such active classrooms necessitate learning opportunities that engage students in a breadth of activities resulting in proof and involving processes which are more inline with how mathematicians engage in constructing proofs (Blanton, Stylianou, & David, 2003; Harel & Sowder, 1998; Smith, 2006; Stylianides, 2007; Stylianides, 2008).

Students' inability to construct proofs stems from the continual handling of proof in classrooms, where instructors present complete solutions with little input from students (Harel & Rabin, 2010; Smith, 2006; Solomon, 2006). This traditional way of teaching mathematics, including proofs, allows instructors to cover more problems and topics including various proof methods. However, when the instructors view their students as passive learners who absorb knowledge and thus do the thinking for their students by presenting complete solutions quickly, students are often left believing that proof construction is an individual instructor activity (Solomon, 2006). The traditional lecture style instructional method leaves students deficient about what is a proof and how to produce proofs on their own, believing they need to memorize what the instructor produces (Moore, 1994; Harel & Rabin, 2010; Solomon, 2006).

A major difficulty with supporting students in constructing proofs is that they do not know what counts as proof. Moore (1994) interviewed 16 mathematics and mathematics education majors including two graduate students and found that “All of them said they had relied on memorizing proofs because they had not understood what a proof is nor how to write one” (p. 264). Without a clear understanding about what is a proof, students continue to view it as a mystery. Stylianides (2007) has proposed and explained his rationale for a definition of proof for school mathematics, which has gained acceptance. He explains seven stipulations to consider after presenting his proof definition. The main point is that classrooms need to co-construct a definition in which the teacher assumes the role of the representative of the greater mathematics community and that his definition can serve as a productive launching point. He proposes that the students and instructor take on active roles during the proving process to continue to develop understanding of the definition of proof and possibly modify it to fit the individual classroom needs, to align with students’ previous conceptions, and to reconstruct understanding of proof. Furthermore, teachers will need to scaffold student thinking as they construct arguments (Blanton, Stylianou, & David, 2003; Stylianides, 2007). The rationale is that as students begin to understand what is needed for an argument to count as proof, through a process that involves reworking their own understandings as conceptual conflicts arise, they will become more successful at constructing them on their own.

Stylianides (2007) proposed a criterion for what counts as proof as follows:

> Proof *is a* mathematical argument*, a connected sequence of assertions for or against a mathematical claim, with the following characteristics:*
> 1. *It uses statements accepted by the classroom community (*set of accepted statements*) that are true and available without further justification;*
> 2. *It employs forms of reasoning (*modes of argumentation*) that are valid and known to, or within the conceptual reach of, the classroom community; and*
> 3. *It is communicated with forms of expression (*modes of argument representation*) that are appropriate and known to, or within the conceptual reach of, the classroom community.* (page 291)

While the criterion is useful to develop a basic understanding of proof, less is known how such a criterion provides supportive feedback to students on the arguments they develop. Such feedback allows student to interact more closely with their previous conceptions of proof as they are drawn into conflict with newer conceptions. In other words a “student friendly” instructional tool is needed for instructors and students to support students’ developing understanding of proof. This article proposes an argument assessment tool that has shown to be applicable to a wide range of possible student solution paths that aligns with Stylianides’s (2007) definition of proof. It is intended to be accessible to students so that they can use it to assess solutions their peers propose and to evaluate their own arguments.

Theoretical Framework

Richardson (2003) calls constructivist pedagogy “the creation of classroom environments, activities, and methods that are grounded in a constructivist theory of learning, with goals that focus on individual students developing deep understandings in the subject matter of interest and habits of mind that aid in future learning” (p. 1627). Based on the idea that students have the capacity to work to build their own knowledge, a constructivist approach to teaching proof necessitates activities that allow students to confront their own, often flawed, understanding of proof and to work towards a more thorough understanding that aligns with the greater mathematics community.

The tool presented in this paper will provide a foundation for constructivist proof instruction. According to Hartle, Baviskar, and Smith (2012), four criteria need to be met for effective constructivism. An activity should elicit prior knowledge, create cognitive dissonance, apply new knowledge with feedback, and reflect on learning.

The Proof Assessment Tool

Many assessment tools to evaluate arguments have been developed for research purposes. Such tools typically include common argument types that are invalid. The research instruments generally range from informal to formal argumentation following some type of numerical or variable scale. For example, Bell (1976) listed two groups with six scales in each one for empirical arguments and one for those that are deductive. The titles include names such as "partially systematic" in the empirical scale and "relevant" in the deductive category. Another approach has been to categorize student solutions into proof schemes (Harel & Sowder, 1998). Harel and colleagues (Harel & Sowder, 1998; Martin & Harel, 1989) evaluated student-produced arguments into proof frames and proof schemes, which both follow an approach, similar to Bell's (1976), along the lines of inductive and deductive reasoning. Finally, Weber and Alcock (2004) approached the categorization from a different perspective; they focused on reasoning that leads to proof production and labeled the categories as syntactic versus semantic. While these approaches to categorizing arguments are useful for research purposes, less is known about how such categorizations are useful to students' development of proof construction. The argument assessment tool (as shown in table 1) aims to fill this gap to better support students.

The argument assessment tool is intended for use in classrooms to develop a shared understanding of proof. That is, it should be used as a starting point from which the classroom community can work to establish a common understanding of proof and develop a common language to discuss a variety of arguments. Because it is to be used to help classes develop shared understanding, it can be used in a variety of settings including secondary math classes, courses for pre-service teachers, and undergraduate proof-based mathematics courses. The argument codes used by Stylianides and Stylianides (2009) are listed in the first column. The code details are added to support instructors and students with more specific feedback as to why an argument aligns with a specific argument code. The final column includes directions on how to code a solution including both an argument code (i.e. A0-A4) and a specific number to provide code details. Three clear and convincing categories are listed below the table and are only applied to valid arguments (codes A3 or A4).

The argument code A0 is used on solutions that are incoherent or do not address the problem situation. A1 is used if the participant was unable to reach a solution or make a generalization instead using only empirical evidence. The code A2 is applied to solutions where the participant attempts a general argument, but is unsuccessful. The various sub codes in the second column (code details) identify specific issues as to why the argument is invalid. The A3 code represents a valid argument that falls short of being a proof. A3 is applied when unjustified assumptions were detected in the argument including assumptions about the conjecture the participant is attempting to prove. It was also applied when the participant lacked the formality required of mathematical proofs within the specific context of classroom community. A4 is applied to proofs, and no sub codes follow because there are no flaws or issues identified. The feedback provided and discussions in classroom communities aim to support the development of a common understanding of the codes. Additionally, what needs to be justified at the beginning of a semester may be left unjustified later as general skill level and understanding progresses.

Since there are differences among valid arguments, a set of clear and convincing criteria are included to provide more detailed feedback to improve arguments. A set of plus or minus symbols are used to code all A3 and A4 main codes. For each of three clear and/or convincing statements, a plus is listed if the criterion is represented, and a minus is used to indicate that a criterion is absent. Therefore, each valid argument code (A3 or A4) is followed by a combination of three plus or minus symbols. For example, a proof that includes unnecessary information and lacks a concluding statement is coded as A4 + - -.

Argument Codes	Code Details	Code Evidence
Incoherent or not addressing the stated problem (A0)	1. Solution shows a misunderstanding of the mathematical content. 2. Ignores the question completely. 3. Interprets claim, provides no argument.	• List A0 and either 1, 2, or 3.
Empirical (example based) (A1)	1. Examples are used to find a pattern, but a generalization is not reached. 2. Only examples are generated as a complete solution. Examples alone are not sufficient! They can be used as a starting point, however.	• List A1 and either 1 or 2
Unsuccessful attempt at a general argument (A2)	1. There is a major mathematical error 2. Illogical reasoning; several holes and or errors exist causing an unclear or inaccurate argument. 3. Reaches a generalization from examples, but does not justify why it is true for all cases. 4. Solution fails to covers all cases. 5. Solution is incomplete. Argument stops short of generalizing the stated claim. Must explain for all cases in general terms; general explanation & all cases covered; logical arguments	• List A2 and match the bulleted number (1-5) in the middle column with the work in the solution.
Valid argument but not a proof (A3)	1. The solution assumes claims, in other words the solution exhibits a leap of faith before reaching a conclusion 2. The solution assumes a conjecture or lists a non-mathematical statement as a conjecture. 3. Argument is sound, but does not use mathematical notation and/or language - too informal No unjustified assumptions, but one can refer to ideas that the community has previously proven.	• List A3 and either 1, 2 or 3 & address each of the points below **
Proof (A4)		• List A4 and address each of the three clear and convincing points below. **
** for use with A3 and A4. (+/-) The flow of the argument is coherent since it is supported with a combination of pictures, diagrams, symbols, or language to help the reader make sense of the author's thinking. Diagrams are fine as long as they are accompanied by an explanation. Explanation of ideas or patterns. (+/-) There are no irrelevant or distracting points. Variables and definitions are clearly defined and any terms introduced by the author are explained. Common understood language (+/-) The conclusion is clearly stated.		

Table 1. Argument Assessment Tool (AAT)

Argument Assessment Tool Used in Two Studies

Because this tool relies on community-accepted norms, the researchers each met with colleagues to discuss what knowledge is accepted in each individual community and what kinds of assumptions need to be justified. For example, due to the different experience levels of the study participants, students in the first study were expected to provide justification for the assumption that the sum of two even numbers is even, but students in the second study were allowed to assume that knowledge without need for justification. After discussion about community norms and an explanation of the assessment tool secondary coders were asked to apply the tool to a variety of arguments. The tool was used to assess arguments in two different studies.

In the first study, 9 graduate students, who all earned an undergraduate degree in mathematics, participated in a reasoning and proving course to prepare them as secondary mathematics teachers in a education department at a large US public university. The first author was the researcher of this study. A second coder used the tool to assess 18 of the 71 student produced arguments. Agreement was reached on 13 of 18 (72%) main codes. Every argument coded as a valid argument (A3 or A4) was also labeled as such by the second coder. There was one instance in which the second coder labeled a solution as a valid argument and the researcher coded it A2. Therefore, four of the disagreements were between A3 and A4 and only one between A3 and A2. There were eight solutions that required sub codes and there was agreement on seven of the eight possible sub codes. There were 33 opportunities to include a plus or a minus for the 11 valid arguments, and 26 of the 33 (79%) instances were agreed upon during the initial coding session. The disagreements were resolved through discussion until agreement was reached on every code.

In the second study, 22 undergraduate mathematics and mathematics education majors were assessed on their proof construction abilities at the start and conclusion of a semester during which the students were enrolled in a variety of proof-based math courses at a large public institution. The PI, the second author, initially assessed the participant-generated arguments with a tool she developed based on validation literature (Selden & Selden, 2003; Weber, 2010) but was frustrated by the coarseness of the tool and the inability to discuss improvement between invalid attempts at general arguments. Using an earlier version of the tool presented here, she was able to assess 130 of the 132 proof attempts, but used the two remaining arguments to refine the tool to the version presented in this paper. The researcher trained two additional coders who looked at 6 of the 84 solutions. All three coders were in agreement on all six main codes, and on only two solutions was there initial disagreement on sub codes. In those instances, after a short discussion agreement was reached on sub codes as well as the clear and convincing ratings.

Coding Examples

Two examples are provided to illustrate how an argument was coded (as shown in Figures 1 and 2). The participants in the first study were asked to prove that $n^2 + n$ is always even for all counting numbers. Tanya's solution (Figure 1) is mathematically correct, but she included an assumption when she wrote "an even times an odd is even" without justifying why this is true. The valid argument with the assumption was coded A3.1. Since the solution is a valid argument, all three clear and convincing statements are checked. The argument does not include jarring statements, missed defined terms or variables, but there is no clear conclusion. Two cases (odd and even counting numbers) are addressed without summarizing the combined argument to explain why the conjecture is indeed true for all counting numbers. Therefore, Tanya's solution was coded A3.1++-. Tanya may not have known that her assumption required justification, but the classroom community agreed on the criterion that any claim that were not previously proven in the course needs to be justified.

Prove that for every counting number n (1, 2, 3, 4 ...), the expression $n^2 + n$ will always be even.[2]

$n^2+n = n(n+1)$

Let n be a counting number.

Then $n^2+n = n(n+1)$.

If n is even, then n+1 is odd.

So n(n+1) is an even times an odd,

so it is even.

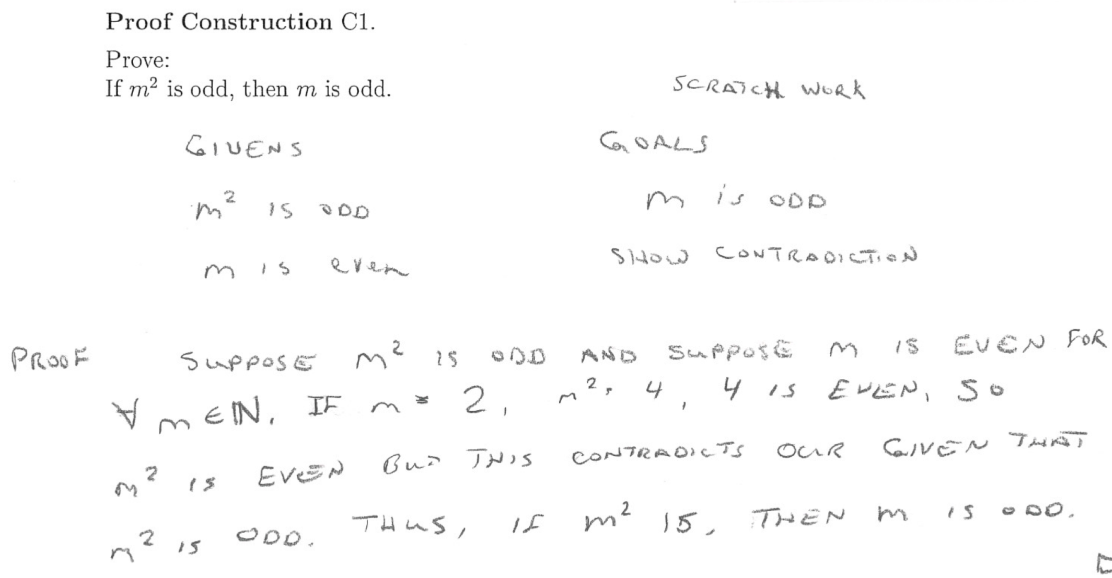
If n is odd, then n+1 is even.

So n(n+1) is an odd times an even,

so it is even.

Figure 1. Tanya's Solution (first study)

The participants in the second study were asked to prove that when m^2 is odd, m is also necessarily odd. Ursula attempted a general argument by contradiction (see Figure 2), but had logical errors that invalidated her argument. She sets up the contradiction argument incorrectly, and contradicted her incorrect statement. A valid contradiction argument would have assumed that m^2 was odd but that the specific m used was even. Assuming that m is even for all m in the natural numbers is a logical mistake that demonstrates a misunderstanding of contradiction. After making that logical error, Ursula correctly contradicted the incorrect statement since a single counter example suffices to negate a for-all statement. There are no other errors present. While the support of her argument is empirical, she was trying to make general statements about all m, so this argument did not fit that A1 code criteria. As such, her argument was rated as A2.2.


Proof Construction C1.

Prove:
If m^2 is odd, then m is odd.

SCRATCH WORK

GIVENS

m^2 IS ODD

m IS EVEN

GOALS

m IS ODD

SHOW CONTRADICTION

PROOF SUPPOSE m^2 IS ODD AND SUPPOSE m IS EVEN FOR $\forall m \in \mathbb{N}$. IF $m = 2$, $m^2 = 4$, 4 IS EVEN, SO m^2 IS EVEN BUT THIS CONTRADICTS OUR GIVEN THAT m^2 IS ODD. THUS, IF m^2 IS, THEN m IS ODD. $\square$

Figure 2. Ursula's Solution (second study)

Conclusion: Future Work

The argument assessment tool was developed because of a need to analyze and provide feedback on student produced arguments. The need for such a tool stems from the instructional goal of moving beyond instructor presented arguments toward supporting students with constructing, evaluating and developing a greater understanding of mathematical proof. Current work with analyzing student arguments has only been from research purposes. While the use of the argument assessment tool has also been used as a research instrument, both authors believe in the potential of the tool as an instructional support to develop students' knowledge of proof in the classroom in a variety of settings.

Both authors used the tool to analyze their participants' arguments, but they were not able to use the tool to provide feedback to the students. The participants in the first study utilized a different set of classroom community constructed criteria of proof to provide feedback to one another during instruction. The argument assessment tool was developed ad hoc as a way to distinguish among the participants' arguments at the end of the course in conjunction with the first study. The participants in the second study were encouraged to talk to the PI at the completion of the study to discuss their arguments, but no participants did. As such, there was not an opportunity to use the tool to give students feedback on their proof attempts. Therefore, while the tool has shown to be a productive assessment tool in different settings with different researchers, the authors do not have data yet on the tool's effectiveness as an instructional resource beyond summative assessment.

Future work will include using it with students in classroom settings as a constructivist activity. Each time students are asked to produce proofs they will necessarily be drawing on prior mathematical knowledge as well as their current understanding of proof production. The authors have plans to use the tool in courses for pre-service teachers as well as courses for undergraduate mathematics majors. Sharing the tool with students openly will allow them to give and receive feedback about proof construction that will often draw attention to the students' misunderstanding of proof, creating cognitive dissonance they will need to resolve. As the students reflect on such feedback individually and as a class, they will be able to revise and redefine the tool as needed to improve their own learning. While this tool is designed for use in a constructivist classroom, the authors hypothesize that it would be of use for grading proofs and providing feedback in other settings as well.

Therefore, future work in classrooms that implement the tool will include study of how students learn to appropriately apply the tool to analyze arguments including their own, construct proofs, and expand their generalized understanding as to what counts and what does not count as proof.

References

Ball, D., Hoyles, C., Jahnke, H., & Movshovitz-Hadar, N. (August, 2002). *The teaching of proof.* Paper presented at the International Congress of Mathematicians, Beijing, China.

Bell, A. (1976). A study of pupils' proof – explanations in mathematical situations. *Educational Studies in Mathematics*, 7, 23-40.

Blanton, M.L., Stylianou, D.A. & David, M.M. (2003). The nature of scaffolding in undergraduate students' transition to mathematical proof. *In the Proceedings of the 27th Annual Meeting for the International Group for the Psychology of Mathematics Education*, v.2, (pp.113-120). Honolulu, Hawaii: Center for Research and Development Group, University of Hawaii.

Edwards, B.S., & Ward, M.B. (2004). Surprises from mathematics education research: Student (mis)use of mathematics definitions. *The American Mathematical Monthly,* 111(5), 411-424.

Harel G. & Rabin J.M., (2010) Teaching Practices Associated with the Authoritative Proof Scheme. *Journal for Research in Mathematics Education,* 41(1), 14-19.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), Research in collegiate mathematics education, III (pp. 234-283). Washington DC: Mathematical Association of America.

Hartle, R. Todd, Baviskar, S. and Smith, R. (2012). A field guide to constructivism in the college science classroom: four essential criteria and a guide to their usage. *Journal of College Biology Teaching*, 38(2), 31-35.

Martin, G.W. & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education,* 20(1), 41-51.

Moore, R.C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics,* 27, 249-266.

Morris, A.K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction,* 20(1), 79-118.

Recio A.M., & Godino J.D. (2001). Institutional and personal meaning of mathematical proof. *Educational Studies in Mathematics,* 48, 83-99.

Richardson, V. 2003. "Constructivist Pedagogy." *Teachers College Record* 105 (9): 1623–1640.

Selden A. & Selden J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education,* 34(1), 4-36.

Smith, J. C. (2006) A sense making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior,* 25, 73-90.

Solomon, Y. (2006). Deficit or difference? The role of students' epistemologies of mathematics in their interactions with proof. *Educational Studies in Mathematics, 61*, 373-393.

Stylianides, A.J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education,* 38, 289-321.

Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics. 72*(2), 237-253.

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 28*(1), 9-16.

Weber, K. (2010). Mathematics Majors' Perceptions of Conviction, Validity, and Proof. *Mathematical Thinking and Learning*, 12(4), 306-336. doi:10.1080/10986065.2010.495468

Weber, K. & Alcock, L. (2004). Semantic and Syntactic Proof Productions. *Educational Studies in Mathematics* 56, 209-234.

Yilmaz, K. (2008). Constructivism: Its Theoretical Underpinnings, Variations, and Implications for Classroom Instruction. *Educational Horizons*, 86(3), 161-172

EXPLORING STUDENTS' QUESTIONS FROM ONLINE VIDEO LECTURES

Fabiana Cardetti, Konstantina Christodoulopoulou, and Steven Pon
University of Connecticut

This study was designed to investigate the types of questions college students generate as they watch video lectures in a business calculus class. Thirty-six students taking an undergraduate calculus course participated in the study. In this paper we share the preliminary results of our qualitative analysis. We have found nine mutually exclusive categories that uncover the thoughts, struggles, and successes our students go through as they experience this new teaching modality of video-viewing. We also include three questions for the audience to help further our analysis and open up new research opportunities for the improvement of collegiate teaching through the study of students' questions.

Key words: Students' Questions, Calculus, Flipped Classrooms, Video Lectures

The exchange of questions between student and teacher forms an integral part of the learning process. There is a wealth of studies devoted to the role of teacher questions and student responses in examining student learning and the effectiveness of teaching approaches. However, we can also learn a great deal from the questions that students ask, since posing questions is a crucial component of learning and scientific inquiry and can play an important role in directing a teacher's practice. The increase in the use of some technologies, especially asynchronous technologies such as video lectures, email, and discussion boards, both heightens the importance of the pedagogical response to student questions as well as allows for more detailed study of student questions.

Conceptual Framework

In their comprehensive review of the research on students' questions, Chin and Osborne (2008) demonstrated that there is considerable educational potential for science teaching and learning in student-generated questions. Students' questions play a key role in science classes directing student learning and driving knowledge construction (e.g. Chin & Brown, 2000), fostering classroom discussion (Chin, Brown & Bruce, 2002), helping students self-evaluate and monitor their understanding (e.g. Chin, 2006; King, 1989), and increasing students' interest in a topic (Chin & Kayalvinzi, 2005). From a teaching standpoint, students' questions can be used to assess learning (e.g. Maskill & Pedrosa de Jesus, 1997; Watts & Alsop, 1995), to evaluate students' higher-order thinking skills (Dori & Herscovitz, 1999), to stimulate scientific inquiry (Crawford, Kelly, & Brown, 2000; Maskill & Pedrosa de Jesus, 1997), and to prompt teachers to reflect critically on their teaching (Watts, Alsop, Gould, & Walsh, 1997).

A key to the study of student questions is their classification. Different classification schemes have been developed to reveal the level of cognitive processes that students use when posing questions in general science classes (e.g. Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Pedrosa de Jesus, Teixeira-Dias, & Watts, 2003; Pizzini & Shepardson, 1991). Such classifications help not only with quantifying and furthering our understanding of student questions, but also with supporting students with a relevant taxonomy that can improve the quality of their questions (Marbach-Ad & Sokolove, 2000).

With regards to mathematics instruction, researchers have emphasized the importance of student-generated questions in mathematics teaching at the K-12 school level and have investigated ways to incorporate them in their classrooms (e.g. Foster 2011; Piccolo, Harbaugh, Carter, Capraro, & Capraro, 2008). However, less has been done to further study student questions at the college level, especially as they relate to the unique features of mathematics, which is more abstract and less experimentally-based than other sciences. With this background in mind, the central research question guiding this study was: What types of questions do college students raise as they watch video lectures in a business calculus class? In addition, we are interested in the relationships between the questions students ask and students' performance related to the videos in the immediate-term (video), short-term (quizzes), and long-term (exam). The research reported in this paper is part of our ongoing research on the appropriate use of technology in undergraduate mathematics (Cardetti, Pon, & Christodoulopoulou, 2013).

Methodology

This study took place in two sections of a college business calculus course in Fall 2012 taught by two experienced instructors. For one particular week in the semester the classes were flipped; that is, students viewed instructor-generated videos prior to the corresponding class meeting and class time was devoted to work on problems from the textbook. After watching the videos, students were required to answer conceptual questions and generate thoughtful questions about the material in the videos. This activity was intended to help verify that students had watched the videos and to encourage them to think critically about the material. The flipped week featured material on finding absolute extrema of functions and real-world optimization problems. Participants included 36 students, most of them were at the sophomore level and about half of them were Business majors.

The primary data source for this study were the questions students generated after watching each of the three sets of lecture videos prepared by their instructors. In total we collected 175 questions. The analysis began with a process of open coding (Strauss & Corbin, 1998) focusing on the general nature of students' questions. We used primarily the constant comparative method to identify emergent categories, distinguishing themes and features of the students' questions. All questions were carefully coded using the initial categories, keeping an eye out for new or more refined versions of the categories to emerge. Each question was independently coded by at least two members of the research team. This was followed by discussions to ensure agreement on classifications and the development of new categories. What we report in this paper are the initial categories that emerged from this data.

We also collected students' answers to the conceptual questions and students' scores on quizzes and exams. The analysis of these data is ongoing and will help find patterns or trends associating students' questions with their performance throughout the term.

Early Results and Development

As our study is ongoing, we will present here findings primarily addressing our main research question regarding student-generated questions. We identified nine broad categories defined below. We also include an illustrative example and a brief rationale for the coding choice. To better understand the coding, it is useful to know that on the videos the instructors presented the standard procedure to find absolute and relative extrema along with examples, as well as optimization problems. In all videos, graphs were used to aid the explanations.

Fundamental Misconception: question reveals basic deficiencies from earlier courses.

"In the second video, where did $S = 2x^2 + 4xh$ come from" (Day 2)

(Student could not follow derivation of surface area of a box.)

Previous material: question reveals lack of recall or relation to material from earlier in the term.

"Why is $\lim_{x\to\infty}(3 - x - x^2) = \lim(-x^2)$ [sic], what is this concept?" (Day 1)

(Finding the end behavior of a polynomial like this was covered previously in the course.)

Basic steps: question reveals difficulty following from one step to the next in the video.

"If there are endpoints that can be possible absolute max/mins how do we decide which one is [the] absolute [extrema]?" (Day1)
(Student seemed to have trouble with the last step in finding absolute extrema on closed intervals.)
Struggling: question reveals student difficulty connecting concepts within the video.
"Why the odd polynomials don't have absolute extrema?" (Day 1)
(Student had trouble connecting a graph of an odd polynomial to the idea of absolute extrema.)
Anxiety: question concerns anxiety over exams.
"Will we ever have a problem like example 2 on a test?"(Day 2)
Mathematical disposition: question reveals students attitude toward mathematics practice or the value of mathematics.
"So if all optimization problems are different, [why] should we study the procedure?" (Day 3)
(Student questions the value of the procedures explained in the video.)
Word problems: question reveals issues with mathematical modeling.
"How can we tell from the beginning of the problem whether or not the objective function will be expressed in terms of one or multiple variables" (Day 3)
Making sense: question centers on a student's first reaction – may be naïve, ingenuous, unsophisticated, or show confusion over a minor mistake.
"*Why* does the process change when using () as opposed to []" (emphasis added, Day1)
(Student seemed interested in further understanding *why* the difference makes sense.)
Non-standard questions: question showcases student thinking beyond what is being presented.
"If there is a hole or jump in the function, does it change anything as to how the absolute extrema are found?" (Day 1)
(Videos dealt only with continuous functions, and the student is exploring degenerate cases.)

Further analysis of our findings will help us refine these categories, as well as find out the frequencies of the different types of questions and whether there are any connections between short- or long-term students' difficulties and the different categories.

Audience Discussion Questions

Our preliminary data analysis suggests a wide variety of questions in students' minds as they watch video lectures. While many of these categories are familiar to the experienced instructor, this study allowed us to survey *all* students in the classrooms, rather than only those who would normally raise questions in a typical in-class lecture. The preliminary results seem to support the importance, raised by the literature, of studying these questions. Moreover, the preliminary findings indicate that the way in which the questions were elicited has great potential to help improve the students' learning experience by requiring them to do in-the-moment deep thinking about the presentation. This suggests further research focused on the nature of this task.

Considering that a better understanding of our students' thoughts and difficulties can play a significant role in informing changes in instruction, there are several questions we would like to discuss with the audience to help further our analysis and direct new research opportunities:

1) How do you see these studies helping your teaching and student learning?
2) How do you see the role of student questions changing in an online or other nontraditional environments?

3) How could we further leverage student questions to get students into the mathematician mindset, allowing them to feel smart and capable of asking good mathematical questions?

References

Bloom, B.S., Engelhart, M.B., Furst, E.J., Hill, W.H., and Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals (Handbook 1: Cognitive domain)*. New York: Longmans Green.

Cardetti, F., Pon, S., and Christodoulopoulou, K. (2013). Flipped Classrooms in College: Calculus Student's Perceptions and Performance. In Gómez Chova, L., López Martínez, A., and Candel torres, I. (Eds.), *Proceedings to EDULEARN13 International Conference on Education and New Learning Technologies*, Barcelona, Spain, 5793–5800.

Chin, C. (2006). Using self-questioning to promote pupils' process skills thinking. *School Science Review, 87*(321), 113–122.

Chin, C. and Brown, D.E. (2000). Learning deeply in science: An analysis and reintegration of deep approaches in two case studies of grade 8 students. *Research in Science Education, 30*(2), 109–138.

Chin, C., Brown, D.E., and Bruce, B.C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education, 24*(5), 521–549.

Chin, C. and Kayalvizhi, G. (2005). What do pupils think of open science investigations? A study of Singaporean primary 6 pupils. *Educational Research, 47*(1), 107–126.

Chin, C. and Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education, 44*(1), 1–39.

Crawford, T., Kelly, G.J., and Brown, C. (2000). Ways of knowing beyond facts and laws of science: An ethnographic investigation of student engagement in scientific practices. *Journal of Research in Science Teaching, 37*(3), 237–258.

Dori, Y.J. and Herscovitz, O. (1999). Question-posing capability as an alternative evaluation method: Analysis of an environmental case study. *Journal of Research in Science Teaching, 36,* 411–430.

Foster, C. (2011). Student-generated questions in mathematics teaching. *Mathematics Teaching, 105*(1), 26 – 31.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14*, 366–381.

Marbach-Ad, G. and Sokolove, P.G. (2000). Can undergraduate biology students learn to ask higher level questions? *Journal of Research in Science Teaching, 37*(8), 854–870.

Maskill, R. and Pedrosa De Jesus, H. (1997). Pupils' questions, alternative frameworks and the design of science teaching. *International Journal of Science Education, 19*, 781–799.

Pedrosa de Jesus, H., Teixeira-Dias, J.J.C., and Watts, M. (2003). Questions of chemistry. *International Journal of Science Education, 25*(8), 1015–1034.

Piccolo, D. L., Harbaugh, A. P., Carter, T. A., Capraro, M. M., and Capraro, R. M. (2008). Quality of instruction: Examining discourse in middle school mathematics instruction. *Journal of Advanced Academics, 19*(3), 376 – 410.

Pizzini, E.L. and Shepardson, D.P. (1991). Student questioning in the presence of the teacher during problem solving in science. *School Science and Mathematics, 91*, 348–352.

Strauss, A. and Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage Publications. 124–125.

Watts, M. and Alsop, S. (1995). Questioning and conceptual understanding: The quality of pupils' questions in science. *School Science Review*, *76*(277), 91–95.

Watts, M., Alsop, S., Gould, G., and Walsh, A. (1997). Prompting teachers' constructive reflection: Pupils' questions as critical incidents. *International Journal of Science Education, 19,* 1025–1037.

MATHEMATICAL PERCEPTIONS AND PROBLEM SOLVING OF FIRST YEAR DEVELOPMENTAL MATHEMATICS STUDENTS IN A FOUR-YEAR INSTITUTION

Anne Cawley
University of Michigan

I report initial findings of a study that seeks to investigate the change in developmental (remedial) mathematics students' mathematical problem solving skills. I report on the analysis of one-on-one interviews with six students before a four-week Intermediate Algebra course. The ultimate goal is to see the extent to which their skills changed after the course. Using a framework of reasoning developed by Lithner (2000), I describe events in which one particular student shows plausible reasoning and also reasoning based on established experience. I seek input with regard to alternative frameworks or analysis of the data that may help me interpret the findings.

Key words: Developmental Mathematics, Problem Solving, Reasoning in Mathematical Tasks

Much research has looked into developmental mathematics students at the postsecondary level (Bahr, 2008; Bahr, 2010; Bonham & Boylan, 2011; Stage & Kloosterman, 1995). This study aims to draw connections between students' prior knowledge with regard to their beliefs of preparedness, self-efficacy and the locus of control (internal versus external) of their mathematical experiences while looking at the change in their mathematical problem solving skills (cognitive strategies). This analysis is especially important in undergraduate education given that more than one out of five college students are required to take remedial mathematics courses varying across arithmetic, pre-algebra, beginning algebra, and intermediate algebra (NCES, 2003). What is more, a majority of students enrolled in developmental mathematics courses need more than one attempt to pass (Attewell, Lavin, Dominia, & Levey, 2006). Postsecondary developmental mathematics education is comprised of courses that include content that is prerequisite for college level courses. The main goal of developmental mathematics courses is to prepare students who are deemed unprepared for college level mathematics courses and bridge them to college level mathematics (Blum, 2007). Stuart (2009) claims that providing help for specific content areas in which they are lacking is more valuable than simply teaching the topics of an entire course over again. Therefore, better understanding the connection of developmental students' misunderstandings or lack of understanding, knowledge of their beliefs as well as drawing connections to their problem solving/cognitive strategies may be essential in finding ways to improve the success of these students. As part of a larger study, the following research question is the focus of this paper: How does developmental mathematics students' mathematical problem solving change after taking a course designed to remediate students' mathematical skills?

Theoretical Framework

Lithner (2000, 2004) proposes a framework of reasoning mathematical tasks via a four-step structure: (1) a problematic situation occurs (where it is not obvious how to proceed), (2) strategy choice (try to choose a strategy that can solve the difficulty), (3) strategy implementation (did the strategy solve the difficulty?), and (4) conclusion (a result is obtained). Lithner describes reasoning as "the line of thought, the way of thinking, adopted to produce assertions and reach a conclusion" (Lithner, 2000, p. 166). What is more, Lithner further describes two different types of reasoning: plausible reasoning and reasoning based on established experiences. Plausible reasoning occurs when the argumentation is founded on mathematical properties and is meant to guide towards what probably is the truth, without necessarily being complete or correct. Reasoning based on established experiences (or superficial reasoning) concerns the transfer of properties from one familiar situation to another situation that has at least superficial resemblance to the familiar situation. I propose to use this framework for this study.

Methods

Six participants responded to recruitment emails to participate in the study. Participants were enrolled in one of two sections of a four-week, developmental Intermediate Algebra course at a diverse, four-year, large, public university located on the West Coast. Students that are required to take developmental mathematics must take the course during the summer before their freshman year at the university. Students in the course who show clear understanding of the material advance forward to college level mathematics in the fall semester. The course is structured for technical majors (e.g., engineering, business) and students that advance from this course may move to pre-calculus algebra, trigonometry, statistics or business calculus, depending on their choice of major.

There are multiple primary sources of data collected over a seven-month time period. Audio recordings of two interviews, fieldnotes from classroom observations, results from a diagnostic mathematics test, final exams, a survey constructed from portions of the Motivated Strategies for Learning Questionnaire (Pintrich, 1991), and a mathematical problem solving task on linear equations were collected. Participants worked independently on a routine problem set with the researcher with the aim to assess their knowledge of linear equations. The problem solving session was video and audio recorded and then transcribed. Each problem solving session lasted approximately 30 minutes in total. The participants were asked to fill in a table of values given a specific set of information. This table includes: (1) graph of the line, (2) symbolic expression, (3) y-intercept, (4) x-intercept, (5) a point on the line, (6) a point *not* on the line, (7) increasing/decreasing, and (8) slope. The problem that Carter, one of the participants, chose to start with included the x-intercept of (0,0), a point on the line (1, -3) and a point not on the line (2,6). The problems are coded in two steps: first I identify "reasoning" situations, then each reasoning situation is coded as either plausible reasoning or reasoning by established experiences. This paper analyzes the first problem from the problem solving session of one particular student, Carter (pseudonym). I chose to focus on his work because out of all participants, he was the most confident student in his interviews and in the classroom observations, yet showed the most varied responses during the one-on-one interview solving the mathematical task.

Findings

Carter demonstrated having experience working with linear equations because he appeared comfortable and confident providing the missing information. There were eight situations that were coded as "reasoning" about what he is doing; only two of these are coded as plausible reasoning. The remaining six situations showcase reasoning by established experience.

Plausible Reasoning

Carter begins the problem with "So far I'm seeing these two [points] will help me find slope here and [be] able to put everything else." He then recalls the slope formula to find that the slope of the line is "-3 over 1 or just simply, -3".

R: So, you used these two points [(0, 0) and (1, -3)] and not this one? [(2, 6)]?

Carter: Because this one's an x-intercept and this one is a point on the line. This one is not a point on the line so if I were to use this one it would give me a different graph. It wouldn't be a graph that… how do I say. I just really follow the words, that's it. That's not on the graph means that it's not a solution to the graph [points to (2,6)]. And these two will have solutions [points to (0,0) and (1,-3)].

Carter reasons, in a mathematical way, what the repercussions are of finding the slope using a point not on the line. While Carter does not state clearly that the "x-intercept" and the "point on the line" are solutions to the line, he does differentiate that the third point will not provide the appropriate slope to the line that was being asked for. Another example that showcases plausible reasoning is when Carter graphs the x-intercept as the horizontal line $y = 0$ and then realizes that the given information indicates that the x-intercept is simply a point. He realizes that the x-intercept should actually be the point (0,0) and not a line. In this case, he noted a contradiction in his reasoning and the given the information. However, when asked about the line he had graphed for the y-intercept, he did not use the idea that the x-intercept is a point; the y-intercept was a line.

Reasoning by Established Experience

Carter mainly reasons by alluding to information he learned previously about linear equations. Twice in the problem, he stated that he could answer something, but then would start writing something different that did not draw the conclusion he claimed to find:

R: Ok, great. What else in that row can you fill in?

Carter: I can definitely do the increasing or decreasing. [writes $y-0=-3(x-(-3))$] […]

R: Ok, so when you did all of this work, what were you trying to answer out of all of these? [points to the top row of missing information]

Carter: One thing…that I did definitely answer is the y-intercept.

Carter seems to recall, in many instances, information that is somehow related to linear equations (e.g., increasing, decreasing), but does not seem to gauge its appropriateness for solving the problem. For example, to answer whether the line is increasing or decreasing, he decides to recall and use the point-slope formula. Ultimately, he does not provide an answer to his original goal and instead veers off to find the y-intercept, of which he also does not use the equation he just derived. Similarly, he does not realize that the equation he finds is indeed one of the missing values that he needs to write into the table. That is, he attempts to consider another, more familiar situation from his past by using the point-slope formula in an effort to aid himself in finding the missing information, without seeing the connections.

Discussion

Analysis of one problem shows that Carter's reasoning within the mathematical task is heavily based upon reasoning by established experience. He uses plausible reasoning to find the slope of the line, but when answering other questions (e.g., finding the y-intercept) he brings mathematical information he recalls, but does not use it appropriately. These findings support the claims made by Lithner (2000) that established experience dominates most students' reasoning. Lithner also argued that while plausible reasoning does occur, it is often local and dominated by established reasoning. Carter had moments where he showed plausible reasoning through the statements he made, and a few times realized a mistake. However, a majority of his reasoning was indeed dominated by more superficial connections. The following questions will be presented to the audience for discussion:

1. What other frameworks might be helpful to analyze and connect these data?

2. Are there other more general conceptual frameworks that will connect this problem solving to other more affective features of the larger project?

References

Attewell, P., Lavin, D., Dominia, T., & Levey, T. (2006). New evidence on college remediation. *The Journal of Higher Education*, 77(5), 886-924.

Bahr, P.R. (2008) Does Mathematics Remediation Work?: A Comparative Analysis of Academic Attainment among Community College Students. *Research in Higher Education*, 49, 420–450

Bahr, P.R. (2010). Revisiting the efficacy of postsecondary remediation: The moderating effects of depth/breadth of deficiency. *Review of Higher Education, 33*, 177-205.

Blum, D. E., Hunter, W., & Schneck, J. (2007). Getting students through remedial math is a constant struggle, but this college keeps trying. *Chronicle of Higher Education*, *54*(9), 62.

Bonham, B. S., & Boylan, H. R. (2011). Developmental Mathematics: Challenges, Promising Practices, and Recent Initiatives. *Journal of Developmental Education*, *34*(3), 2-4.

Lithner, J. (2000). Mathematical reasoning in task solving. *Educational studies in mathematics*, *41*, 165–190.

Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *The Journal of Mathematical Behavior*, *23*(4), 405-427.

National Center for Education Statistics. (2003). *Remedial Education at Degree-Granting Postsecondary Institutions in Fall 2000* (NCES 2004-010). Washington, DC: Parsad, B., Lewis, L., Greene, B.

Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. *Student perceptions in the classroom*, 149-183.

Stage, F. K., & Kloosterman, P. (1995). Gender, beliefs, and achievement in remedial college-level mathematics. *The Journal of Higher Education*, 294-311.

Stuart, R. (2009). Reinventing Remedial Education. *Diverse Issues in Higher Education, 26*(18), 14-17. Retrieved from http://search.proquest.com.proxy.lib.umich.edu/docview/194193771?accountid=14667

39

ON THE SENSITIVITY OF PROBLEM PHRASING – EXPLORING THE RELIANCE OF STUDENT RESPONSES ON PARTICULAR REPRESENTATIONS OF INFINITE SERIES

Danielle Champney
California Polytechnic State University

This study will demonstrate the ways in which students' ideas about convergence of infinite series are deeply connected to the particular representation of the mathematical content, in ways that are often conflicting and self-contradictory. Specifically, this study explores the different limiting processes that students attend to when presented with five different phrasings of a particular mathematical task - $\sum(1/2)^n$ *- and the ways in which each phrasing of the task brings to light different ideas that were not evident or salient in the other phrasings of the same task. This research suggests that when attempting to gain a more robust understanding of the ways that students extend the ideas of calculus – in this case, limit – one must take care to attend to not only students' reasoning and explanation, but also the implications of the representations chosen to probe students' conceptions, as these representations may mask or alter student responses.*

Keywords: Calculus, Infinite Series, Interviewing, Tasks

Background and Framing

Existing literature on students' understanding of infinite series (i.e. series of numbers, Taylor series, power series) is extremely sparse, despite the overwhelming notion that it represents (1) the most important topic that students can understand from traditional second semester calculus, if they are preparing for a future in engineering, physics, and other related fields (e.g. Alcock & Simpson, 2009; Tall & Schwarzenberger 1978), and (2) the topic with which students have the most difficulty, when considering the entirety of the traditional second semester calculus syllabus (e.g. Monaghan 2001; Biza, Nardi, & Gonzalez-Martín 2008; Champney, in preparation). Literature suggests an unbalanced treatment of the related material, as "sequences are played down, or even omitted, whilst Taylor series, geometric series, series expansions for the exponential, sine, cosine, etc are a fundamental part of sixth form work," (Tall & Schwarzenberger 1978). Alcock and Simpson (2005) find that even when presented with a definition of convergence of an infinite series and asked to paraphrase it, directly following a unit of instruction on the topic, students differ in their descriptions of what it means for a series to converge, with a large percentage of them providing a mathematically incorrect description. Thus, it is not surprising that, as a field, we document a very wide range of ways that students think about "converge" (Monaghan 1991), in its broadest sense.

One of the most common tripping points for students in their study of infinite series comes in making the transition from studying infinite sequences to studying infinite series of numbers. While series of numbers are for mathematicians a natural extension of the study of infinite sequences, students get tangled in the complexities of the many different limiting processes that they must coordinate, as they turn the notion of infinite sequences on its side in order to reframe and accommodate their new ideas of infinite series. Namely, they must make the shift from considering simple limits of sequences to redefining an infinite series as a sequence of partial sums, and attend to the limit of *that* sequence, and not the sequence of terms itself, if they wish to make claims of convergence. Confusion abounds as students struggle with reconciling the different visual representations associated with the sequence of terms vs. sequence of partial sums, etc… and the language of limits gets used and abused as students attempt to explain the meaning of convergence in this new context. Champney (in preparation) explores this idea further in a large-scale study, which identifies bottlenecks and tripping points for students, as they begin to adapt their notions of limit in these new ways.

In order to sort out students' confusion, and get a more robust understanding of what students understand about the phenomenon of convergence of infinite series – and how this is and is not consistent with their more fundamental notions of limit in calculus – it is reasonable to turn to the limit literature and ask students to explore and explain some of the tried and true examples, such as "Explain why $0.\overline{9}$=1" (from Oehrtman, 2002, for example). However, in using this task with the intention to understand what particular aspects of convergence students find salient as they make connections between limits, sequences, and series, one may have equivalently asked "Is $0.\overline{9}$=1? Please explain." or "To what does the sequence {0.9, 0.99, 0.999, 0.999, … } converge?" While mathematicians would see the consistency in these various phrasings of the same question, and reasonably assume that if students have some idea about the nature of the mathematical object $0.\overline{9}$, then they will answer consistently across all phrasings, prior literature and the study discussed in this paper indicate that this is not the case. Consider the "$0.\overline{9}$" task – Figure 1 demonstrates five different ways that this task has been used in the limit literature, dating back several decades, all for different purposes and with different results. The second column in Figure 1 displays analogous phrasings, using $\sum(1/2)^n$, henceforth referred to as the "halving" task (Champney finds that students use the "halving" task more readily than $0.\overline{9}$ when asked for accessible examples of infinite series). One may guess that while students may be able to reason with particular phrasings of either of these questions, other phrasings are presented in ways that do not align with students' understanding of convergence. In the most extreme sense, these alternative phrasings may be presented in such a way that masks students' true understanding of convergence *or* brings different features of the mathematics to light in such a way that they interpret certain phrasings of the question entirely differently than others.

Some work has been done to explore the effects of multiply phrasing tasks such as these. One recent example (in a physics context), Wittmann (2012) demonstrated the differences that resulted in asking students the same task about several bulbs in a circuit by (1) framing the question as asking for both an answer *and* a justification of that answer vs. (2) framing the question by providing the answer and asking *only* for justification of that answer. In his study, Wittmann curiously found that, though many students were not able to choose and justify the correct answer for themselves, those same students were largely able to provide adequate justification for the correct answer, when it was provided.

While it may be interesting to pose the question: to which phrasings of the "halving" or $0.\overline{9}$ task do students respond correctly/consistently with a more formal understanding of the topic of convergence, this is *not* the question explored in this study. Knowing which phrasings are more or less likely to prompt 'correct' responses does not provide one with any information on the range of ideas that students associate with the topic, nor give guidance for improving student understanding. Thus, in this study, I aim to address the following research questions, which are more aligned with exploring students' extensions of their limit understanding, and more directly impact future instruction aimed at helping students grasp the difficult concept of infinite series:

- What aspects of students' understandings of convergence (of infinite series) are illuminated by the different phrasings of the "halving" question, outlined in Figure 1?
- How are these differences significant in the way we calculus educators (a) frame our teaching of this content, and (b) assess students' understanding of this content?

Data and Methods

In 2010, semi-structured interviews designed to investigate students' spontaneously generated visual representations used when explaining the topic of infinite series (of numbers) to a less-knowledgeable peer were conducted with second and third semester calculus students, and real analysis students. It became obvious during the course of the first round of interviewing that most students used some version of the $\sum(1/2)^n$ example on their own, to explain convergence, at some point during the interview. That is, it appeared to be an example with which students were relatively familiar and had some level of comfort in using to support their explanations. Thus, during the flow of each subsequent interview, the questions in the right column of Figure 1 were posed, at various times, as they became relevant and related to students' explanations and considerations through the course of their explanations. The purpose of posing these questions, as discussed above, was to explore which different aspects of convergence were illuminated with a consideration of each different phrasing of the problem.

The study discussed here makes use of a particular 1.5-hour semi-structured interview with sophomore engineering major Jenna – a representative of the larger sample of students, all of whom were participating in the related study on visual representations for infinite series. At the time of the interview, Jenna had successfully completed second semester calculus, which included an extensive unit on infinite series (taught from Stewart), with a grade of "B." Jenna was chosen as a representative case because while her responses were very typical among the larger sample of participants, she communicated them more clearly and thoughtfully than others. An in-depth, microgenetic analysis of this interaction with Jenna takes a very close look at the particular limiting processes and aspects of infinite series that were prompted by her reasoning with each phrasing of the "halving" task, and allows for a more fine-grained level of analysis than would be ascertained in other assessment situations (Calais, 2008). Such analysis with all student interviews would be impractical, so the discussion of Jenna is followed by some general patterns observed from the entire set of student interviews.

The Case of Jenna – in brief

While it is difficult to condense Jenna's work with the various phrasings of the "halving" task to such a confined space, what follows is a short description of her responses to tasks 2.1-2.5. Much more detail and discussion is provided in the more extended analysis. In the following description, of interest (as it pertains to the research questions) are the particular aspects of the "halving" task itself that Jenna considers when making conclusions about convergence. It will become apparent that in each different phrasing, Jenna shifts her attention to a different mathematical structure used in that particular task, which causes differences in the way that she views the phenomenon of convergence of this particular infinite series.

During her teaching episode, Jenna spontaneously brought up examples with both $0.\overline{9}$ and $\sum(1/2)^n$ on her own, as part of her explanation of series convergence. Her use of these examples provided entry points for all of questions 2.1-2.5 to be posed, at various relevant points in her explanation. And while the representations that she used to explain and the conclusions that she drew about the general topic of infinite series were quite consistent, Jenna showed vastly different understandings when responding to the different phrasings of the "halving" task.

In brief, when responding to task 2.1, Jenna was able to provide a complete response for why $\sum(1/2)^n = 1$. She first simply recalled the "formula" for convergent geometric series, and discussed how this fit the model of a geometric series whose common ratio of 1/2 indicated that it converged, in particular to a value of 1. When pressed by the interviewer for more reasoning and understanding, Jenna was able to provide both a "walking to the wall" metaphor (reminiscent of Zeno's Paradox - in which she used halving distances to describe the terms of the series, and the distance from her to the wall of 1 yard) and a geometric representation drawn on the whiteboard to further describe her reasoning (see Figure 2). By this work, one might claim that Jenna has some idea of what it means for an infinite series to converge, calling to attention her correct use of an appropriate metaphor, formula, and visual representation. Her examples were good, and aligned with some of the more traditional ones that might occur in a lecture on the material. She clearly communicated an idea of what a geometric series is, and was able to make conclusions based on it, beyond simply recalling the "formula" for convergent geometric series. Her work on task 2.2 extended this – in Jenna's words, "You *can* compute $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+...$ because even though it goes on forever … that's just a geometric series that converges because *r* is less than 1."

In her response to task 2.2, if stopping there, one might conclude that this is further evidence of her understanding the topic at hand. While it is unclear if Jenna could produce a proof or explanation of the more general case to which she refers – $\sum a(r)^n$ – it is clear that she is able to appropriately discuss the different parts of this general type of series. However, to this point in the interview, her attention has largely been on a term-by-term comparison, emphasizing the action of halving of distances or "pies" (as in Figure 2), and appears to be considering, almost exclusively, the <u>individual terms</u>, and the <u>resulting compilation of terms</u>, rather than any other limiting processes that may be appropriate in this scenario.

Continuing to reason with task 2.2, however, Jenna goes on to say "those later terms get so small that they don't matter, and after a while it doesn't change the sum." This is an inappropriate extension of limit ideas that sheds first light on some ways in which Jenna's understanding of series convergence is not correct. It also marks a shift in her reasoning pattern, away from the individual terms and toward the <u>sequence of terms</u>. That is, her attention has shifted to instead considering the ordered list of individual terms, and not just the terms independently, as she decides to consider that the later terms' magnitude is so small that it must not impact the overall sum. While consideration of independent, individual terms allowed her to justify why $\sum(1/2)^n = 1$ (task 2.1), consideration of the ordered sequence of terms lead her to the same conclusion, but for a different reason, in task 2.2.

Additional contrasting evidence comes in her response to task 2.3, in which Jenna claims that, rather than producing a value of 1 (answer choice (b)), which would align with her earlier responses, the sum $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+...$ is "just less than 1, by some infinitely small value," (answer choice (c)). When probed, Jenna's unexpected answer choice was justified in a way that indicated a further shift in the mathematical structure to which she was attending. The presentation of the infinite series in task 2.3 – as an expansion with ellipses after five terms – caused her to conclude that this series must be "approaching, but not quite 1," unlike task 2.1, which was presented with sigma notation. Thus, a simple shift in the presentation of the infinite series in question caused Jenna, by her own explanation, to arrive at a different conclusion about the convergence of said series. Now, rather than attending to the individual terms or the sequence of ordered terms, Jenna attended to the <u>ellipses</u> as an indication of a different sort of limit process, for which she was uncomfortable concluding that the series converged to 1.

Further complicating matters, Jenna viewed tasks 2.4 and 2.5 as inconsistent with the other questions, and was unable to leverage her understanding to talk about partial sums or the sequence of partial sums. Task 2.5, in particular, was meant to examine whether writing the values as decimals would call something different to mind than the terms expressed as fractions. However, Jenna (as representative of many other students) was fully able to connect those values to values of "partial sums" of $\sum(1/2)^n$ without a problem. Even still, though she had previously used the language of "converges to 1," for tasks 2.4 and 2.5 Jenna said the strongest statement she could make was that it "tends to 1" or "approaches 1." In fact, when posed as a sequence of partial sums, Jenna explicitly denied that one could claim "convergence (to 1)." Jenna's work with the five phrasings of the "halving" task can be summarized (briefly) in Figure 3.

Thus, I claim that the differing responses that Jenna made to the different phrasings of the $\sum(1/2)^n$ task are interesting and useful, but that only when we consider the collection of responses to the different phrasings do we fully understand (a) the full scope of what she really intends as the meaning of "converge," (b) potential ways in which Jenna coordinates the different limit ideas with the different mathematical structures when considering infinite series, and (c) which of these understandings are dependent on task presentation vs. more deeply tied to her knowledge of the content. Thus, looking at her response to any one of these tasks can tell us about *some* aspect of her understanding, that understanding is only tied to the particular representation of $\sum(1/2)^n$ used in that task. But that is insufficient to say that she has a particular "model" (a la Williams, 1991), or "misconception" (a la Davis & Vinner, 1986; Cornu, 1991; and more). Simply asking the questions differently altered Jenna's responses, demonstrating that the different features of a particular task called to mind differences in the way that she interpreted the notion of convergence.

As will be discussed in much greater detail in the full report, though Jenna's responses were representative of the larger sample that participated in this study, there were additional patterns of responses that are significant. For example, there was a large overlap *within student*, consistent with Wittmann's (2012) findings – a significant portion of students were both able to explain why $\sum(1/2)^n = 1$ (task 2.1) while also claiming that one cannot compute $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+\ldots$ and get an answer (task 2.2). Potential explanatory factors for this discrepancy are discussed at length in the full report, but they go beyond Wittmann's suggestion that knowing the answer frees the student to justify and use reasoning, rather than trying to first make a choice and then justify it. The factors that appear to, in part, explain the overlap in this data are more closely tied to the mathematical content and the ways that students orient to infinite series as presented in a variety of different formats – with sigma notation, as a sequence of terms, as a sequence of partial sums, with ellipses v. with a limit symbol, and more. This and other patterns are identified and discussed at more length in the full report.

Contributions and Implications

Often in the present limit literature, it is common to find "results" that attempt to characterize student understanding based on their response to only one framing of, for example, the $0.\overline{9}$ task. While conclusions based on a singular phrasing of a task may be locally relevant to a particular researcher's agenda at hand, the study discussed here indicates that using a singular representation to make some claim about students' limit understanding is inadequate, as the way that the task itself is phrased has significant influence on the particular mathematical structures to which the students attend – thereby influencing the ways that they appeal to their understanding of phenomena such as convergence. From the many ways that the $0.\overline{9}$ or "halving" questions could have been posed, and the many ways that a single student could (and often does) respond differently, depending on the way it was phrased, it seems speculative to claim that any student has a particular understanding of such a concept that is not tied explicitly to the way that the mathematics was represented in the task. Thus, the findings here speak to and have implications for two general audiences. First, for those interested in how students extend their understanding of limit, it important to note how the subtleties in task presentation unintentionally bring to light differing limiting processes that have enormous impact on student interpretations of convergence. Generalizing this, broader audiences can take away not only content implications, but also methodological implications for the specificity of claims about student understanding, as they relate to task presentation.

References

Alcock, L. & Simpson, A. (2005). Convergence of Sequences and Series 2: Interactions Between Non-visual Reasoning and the Learner's Beliefs about their own Role. *Educational Studies in Mathematics, 58,* 77-100

Alcock, L. & Simpson, A. (2009) Ideas from mathematics education: An introduction for mathematicians.

Calais, G. (2008). Microgenetic analysis of learning: Measuring change as it occurs. *National Forum of Applied Educational Research Journal, 21*(3).

Monaghan, J. (1991). Problems with the language of limits. *For the Learning of Mathematics, 11*(3), 20-24.

Monaghan, J. (2001). Young people's ideas of infinity. *Educational Studies in Mathematics, 48* (2&3), 239-257.

Nardi, Biza, & Gonzalez-Martin. (2008). Introducing the concept of infinite sum: Preliminary analyses of curriculum content. In Joubert, M. (Ed.) *Proceedings of the British Society for Research into Learning Mathematics*, 28(3), 84-89.

Oehrtman, M. (2002). Collapsing dimensions, physical limitation, and other student metaphors for limit concepts: An instrumentalist investigation into calculus students' spontaneous reasoning. Doctoral Dissertation. The University of Texas, Austin, TX.

Szydlik, J. (2000). Mathematical beliefs and conceptual understanding of limit of a function. *Journal for Research in Mathematics Education 31*(3), 258-276.

Tall, D. & Schwarzenberger, R. (1978). Conflicts in the learning of real numbers and limits. *Mathematics Teaching, 82*, 44-49.

Tall, D. & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12,* 151-169.

Williams, S. (1991). Models of limit held by college calculus students. *Journal for Research in Mathematical Education*, *22*(3), 219-236

Wittmann, M. (2012). Epistemological framing as a lens on students, teachers, and researchers. Talk presented at the Second Conference on TRUSE (Transforming Research in Undergraduate STEM Education), June 2012, St. Paul, MN.

Figures

(1.1) Q: Explain why $0.\overline{9}=1$ (Oehrtman, 2002)	(2.1) Q: Explain why $\sum_{n=1} \frac{1}{2^n} = 1$
(1.2) Q: Can you add 0.1 + 0.01 + 0.001 + ... (the dots indicate continuation) and get an answer? (Monaghan, 2001)	(2.2) Q: Can you compute $$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+...$$ (the dots indicate continuation) and get an answer?
(1.3) Q: What is between 0.999… and 1? (a) Nothing because 0.999… = 1 (b) An infinitely small distance because 0.999… < 1	(2.3) What is the value of $$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+...?$$ (a) It does not have a value because it keeps

(c) You can't really answer because 0.999… keeps going forever and never finishes. (d) If you agree with one of the above, provide your own answer. (Szydlik, 2000)	going forever and never finishes (b) 1. (c) Just less than 1, by some infinitely small value (d) If you agree with none of the above, provide your own answer
(1.4) Q: Find the limit of the sequence: $\lim_{n\to\infty}\left(1+\frac{9}{10}+\frac{9}{100}+...+\frac{9}{10^n}\right)$ (Tall and Vinner, 1981)	(2.4) Q: Find the limit of the sequence: $\lim_{n\to\infty}\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}...+\frac{1}{2^n}\right)$
(1.5) Q: Consider the sequence {0.9, 0.99, 0.999, 0.999, … }. Which of the following is true of this sequence? (a) It tends to $0.\overline{9}$ (e) it tends to 1 (b) it approaches $0.\overline{9}$ (f) it approaches 1 (c) it converges to $0.\overline{9}$ (g) it converges to 1 (d) its limit is $0.\overline{9}$ (h) its limit is 1 (Monaghan, 1991)	(2.5) Q: Consider the sequence {0.5, 0.75, 0.875, 0.9375, 0.96875, …}. Which of the following is true of this sequence? (a) It tends to 1 (b) It approaches 1 (c) It converges to 1 (d) Its limit is 1 (e) It tends to some value that is not 1 (f) It approaches some value that is not 1 (g) It converges to some value that is not 1 (h) It has a limit that is not 1

Figure 1: Comparison of tasks – existing literature and current study

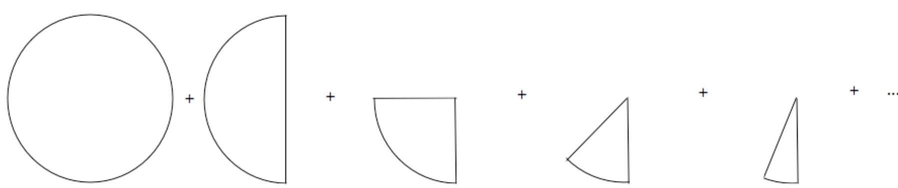


Figure 2: Jenna's geometric representation of the "halving" scenario

(2.1) Q: Explain why $\sum_{n=1} \frac{1}{2^n}=1$	Attended to: individual terms, independently Explanation: included geometric image (see Figure 2), "walking to wall" metaphor, and use of geometric series "formula"
(2.2) Q: Can you compute $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+...$ (the dots indicate continuation) and get an answer?	Attended to: sequence of ordered terms and the decreasing size of "eventual" terms Explanation: if the terms get "small enough" then they become "negligible" and you *can* get an answer
(2.3) What is the value of $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+...?$ (a) It does not have a value because it keeps	Attended to: the ellipses at the end of the series as an indication of continuation and uncertainty Explanation: The "dot dot dot" means "getting very close to, but not reaching"

going forever and never finishes (b) 1. (c) Just less than 1, by some infinitely small value (d) If you agree with none of the above, provide your own answer	
(2.4) Q: Find the limit of the sequence: $$\lim_{n\to\infty}\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}...+\frac{1}{2^n}\right)$$	Attended to: the <u>limit symbol</u> and the idea that there appears to be a "last term" in this version of the task Explanation: the limit is 1, but this version of the task is "not related" to $\sum(1/2)^n$ because this limit is not the same as the $\sum$ symbol
(2.5) Q: Consider the sequence {0.5, 0.75, 0.875, 0.9375, 0.96875, …}. Which of the following is true of this sequence? (a) It tends to 1 (b) It approaches 1 (c) It converges to 1 (d) Its limit is 1 (e) It tends to some value that is not 1 (f) It approaches some value that is not 1 (g) It converges to some value that is not 1 (h) It has a limit that is not 1	Attended to: the sequence as <u>individual values that represent the various partial sums</u>; (When prompted) shifted attention to the collection of these partial sums as an ordered <u>sequence</u> Explanation: the sequence "approaches 1" and "tends to 1" but does not converge to 1, because this "can never equal 1."

Figure 3: Jenna's responses to the five phrasings of the "halving" task

VISUALIZING MATHEMATICAL CONNECTIONS IN STUDENT TEACHING EPISODES

Danielle Champney
California Polytechnic State University

Keywords: Calculus, Infinite Series, Video interview data

This poster aims to present a modified version of SPOT diagrams (Structure Perceived Over Time) (Yoon, 2012) – an aspect of analysis and data presentation used to present interactive student video data, during which perceptual shifts may occur. The larger study in which this tool was employed (Champney, in preparation) explored undergraduate calculus students' self-generated representations (SGR) used during interviews in which they were asked to explain to an absent peer the events of the day(s) during which infinite series were introduced and discussed. While typical studies ask students to address tasks and issues framed by a researcher, this study instead asked students to *explain* the content, thereby providing a broader window into "what counts" from the student perspective.

The main analysis indicated that, while they used the same general image types, the connections drawn between nine identified limiting processes (e.g. sequence of terms, generic compilation, sequence of partial sums) was among the most significant factors separating experts' and students' arguments about convergence of infinite series. While it is important to understand which limiting processes (see also Tall, 1980) students find important when reasoning about infinite series, simply knowing which ones students draw on most frequently is insufficient to understand what sense students make of this topic. It is further necessary to uncover the ways that students coordinate and connect these limiting processes, while explaining. Studying the connections that students make allows insight into which limiting processes are prioritized and why, in a student's attempt to produce a coherent, mathematical story that accounts for infinite series convergence. This is difficult, however, because of the interactive nature of the students' teaching episodes.

The modified SPOT diagrams show how students connect ideas about infinite series similarly and differently, and help to organize and elucidate which mathematical structures students have constructed, prioritized, and found relevant to the context. The use of such diagrams is instrumental in the study at hand because they help to demonstrate how students connect the limiting processes that they find important in supporting their arguments about infinite series convergence, and help to identify whether some learning or perceptual shifts may have actually occurred *during* the teaching episode itself, amid the student's explanation.

SPOT diagrams have been shown to be particularly useful at identifying factors that contribute to "aha! moments" during exchanges in which students' initial ideas represent the potential for 'genuine conceptual development' (Yoon, 2012). Several such teaching episodes were identified, and a modified form of a SPOT diagram is being explored with one student, Molly, who had a literal "aha!" moment during her explanation. Molly's teaching episode is unique because, though she started with an admittedly weak understanding of infinite series ("partial sums is a word *they* use," "I think a graph might be involved," "I think I understand sequences better than series"), she proceeded to spend the following 24 minutes constructing an internally consistent account for the meaning of series convergence. While not always mathematically sound, the connections drawn among limiting processes in her account are both consistent and built around eight SGR that she spontaneously produced. Thus, a modified SPOT diagram provides an empirical lens into shifts and connections in Molly's understanding, which allows for the examination of the events of her teaching episode from multiple perspectives on learning. Using Molly as a single case, the aim of this poster will be to share the affordances of these modified SPOT diagrams as tools for organizing and presenting complicated video data, during which perceptual shifts may have occurred in students' understanding.

References

Tall, D. (1980). Mathematical intuition, with special reference to limiting processes. *Proceedings from the Fourth International Conference for the Psychology of Mathematics Education,* 170-176.

Yoon, C. (2012). Mapping mathematical leaps of insight. Proceedings from the 12th Annual International Congress on Mathematical Education.

What are we sure about? What do they tell about our probabilistic thinking?

In this study the prospective teachers' understanding of extreme probabilities is studied via their examples. Watson and Mason's Learner Generated Examples (LGE) theory is employed to justify the type of data used in this study and to emphasize the importance of examples in learning about different levels of the learners' probabilistic thinking.
Keywords: Probability, Learner Generated Examples, Teacher Knowledge

About the LGE framework:

Watson & Mason (2005) considered Learner generated Examples (LGEs) -an approach in which learners are asked to provide examples of mathematical objects under given constrains – as a powerful pedagogical tool, through which learners enhance their understanding of the concepts involved. Watson and Mason also introduced the construct of example space as collections of examples that fulfill a specific function, and distinguished among several kinds of example spaces. When invited to construct their own examples, learners both extend and enrich their personal example spaces, but also reveal something of the sophistication of their awareness of the concept or technique (Bills, Dreyfus, Mason, Tsamir, Watson, & Zaslavsky, 2006). In accord with this observation, Zazkis and Leikin (2008) suggested that LGEs provide a valuable research tool as they expose learner's ideas related to the objects under construction and examples generated by students mirror their understanding of particular mathematical concepts. Of my interest in this study are personal example spaces, triggered by a task as well as by recent or past experience, and collective example spaces, local to a classroom or other group at a particular time.

Mason in an analysis of the phenomenology of example construction (Mason, 2011) describes what takes place through the process of mathematical example construction as: A strong tendency to combine the simplest possible with maximum generality, constructing lots of examples and tinkering with examples to modify them so that they meet some particular constraint, experiencing dimension of possible variation and range of permissible change associated with the examples constructed and explore deeper aspects of the notion, and drawing attention to the playful aspects of example construction and the ways of tinkering with a basic construction that might be of benefit the future use. Vinner (2011) finds the role of examples in everyday and mathematical thinking to be very crucial. Unlike in mathematics in which the concept formation is aided by definitions, examples and proofs, in everyday thinking, examples are the only tool by which we can form and verify concepts and conjectures. Even in mathematics there are important notions such as "proof" that have no (undergraduate level) definitions and the students are supposed to acquire the concept of proof by the many examples they are exposed to.

Zazkis & Leikin (2008) suggest that The task of constructing examples of mathematical concepts can be quite a complex task for students and teachers, but several researchers find it a well worth effort since the example generating task provides rich educational potentialities: providing a window into learner's mind through which significant aspects of conceptualization could be observed, raising the students' awareness of features of examples that can change and of the range where they can vary (Mason, 2011), and the richness and complexity of processes involved in constructing examples (Antonini, 2011).

Examples also may be used to identify, mirror, and confront learners' incorrect mathematical inferences. Building on ideas of cognitive conflict and conceptual change, Zazkis and Chernoff (2008) extend the considerations of dimension of possible variation and range of permissible change to counterexamples and discuss the role of counterexamples with respect to those theoretical constructs while helping students face their misconceptions.

Many different ways to look at the examples are introduced into the research. From the generating point of view, there are two types of examples: those generated by learners upon invitation (learner generated examples) and the examples used by teachers in a classroom setting (instructional examples). With regard to the availability of examples to the generator one can distinguish between situated, personal, personal potential, accessible, and conventional example spaces; discussed in Watson and Mason 2005. With regard to the specific functioning of examples one can put them into examples-of, examples-for (Michener, cited in Watson 2011), pivotal examples, bridging examples (Zazkis & Chernoff 2008), non-examples, counterexamples, ...

This study:

In this study I consider student generated examples of an event with 100% probability and address the following question:

To what extent do examples generated by participants reveal their understanding of the mathematical concept of probability and more specifically of the certain events?

About 100% probable events:

Extreme probabilities have mathematical significance. Also known as tail probabilities, the extreme probabilities create additional complexity to the probability estimation methods and techniques. Every computer simulation method has limitations and problems when the probability sought after is around the extremes. For example the central limit theorem allows for a binomial distribution to enjoy the normal approximation when np and n(1-p) are both greater than 5, even if the sample size is small. For very extreme probabilities, though, a sample size of 30 or more may still be inadequate and the approximation works at its worst when the sample proportion is exactly zero or exactly one.

From an educational perspective distinguishing between the binary opposites of certain-uncertain and possible-impossible is often located at the very introductory phases of a typical probability education. For example Van de Walle (2011), suggests that young children come to class with all sorts of bewildering ideas of probability, "to change these early misconceptions, a good place to begin is with a focus on possible and not possible and later impossible, possible, certain" (p. 474). Thus extreme probabilities are the type of events that the learners are familiar with since the very early grades.

Participants of the study:

The participants of the study are 29 undergraduate students taking a mathematics education course in Simon Fraser University, Vancouver, with a diverse mathematical background including arts, social studies, biology, and computer science. However, all of them have taken an equivalent of an introductory probability and statistics course at some point before. They are asked to give examples (in writing) of events with 100% probability of happening. They produce 45 examples in total. The task is presented to them in written form and the time for answering has been unlimited.

Method of data analysis:

The Framework used to analyze the data is a tool for analyzing personal or collective example spaces based on (a) correctness, (b) richness, and (c) objective-subjective duality. The first two elements of this framework are adopted form Zazkis & Leikin (2008) the last part is borrowed form Gillies (2000) and Chernoff (2008).

Correctness: In the *correctness* category I consider whether the examples satisfy the condition of the task, which is fulfilling a 100% probability of happening from a reasonably acceptable mathematical point of view. There is an important decision to be made before we go through a discussion of correctness of data. If a student expresses a belief that there is a 100% chance that the next roll of a die will be six and to prove himself correct he rolls a die and a "six" shows up, is his assessment of the probability correct or not? It is while we do not have any knowledge of the die, it could be a fair die or it could be loaded to show six all the time. The same issue comes up in several examples from the participants: "there is a 100% chance that I will take the bus back home today" is this a correct example of a certain event or not? The student that has presented this example possesses certain knowledge of her transportation options and habits and perhaps she sees this as the only event in the sample space and perhaps she is right to assign a 100% probability to it.

The criteria for assessing the validity of the probabilities assigned to events is whether common sense (to be more specific: accounting for all of the possible scenarios/outcomes) and knowledge that is reasonably accessible to everyone is used and wherever applicable the background information necessary to make the judgment appeal to other people is presented or not. For instance in the bus example it is reasonable to take into account that a bus is a vehicle prone to accidents or general failure and it simply might break, so there is a chance however small that she might have to call a friend to give her a ride back home today. That marks this particular example incorrect and I have coded them as *lacking key information or common sense (Lack).*

Another group of examples that have been identified as incorrect are what I call examples of "non-random events". In this group of examples the participant holds a vision of 100% probability as a fact that no one can challenge or refute, and finds such facts from situations which are not subject to randomness at all simply because they either refer to events in the past or they deal with definitions. Among such examples are: "There is a 100% probability of me having the same color eyes as someone in this class because I see some people with the same eye color", "There is a 100% chance that tomorrow is Tuesday given that today is Monday", and "There is a 100% chance that I went to bed before 10 last night, because I did so". This group of examples is coded as "Non-random situations" (NR)

Table 1: correctness of the participants' examples (n=45) for events with 100% probability

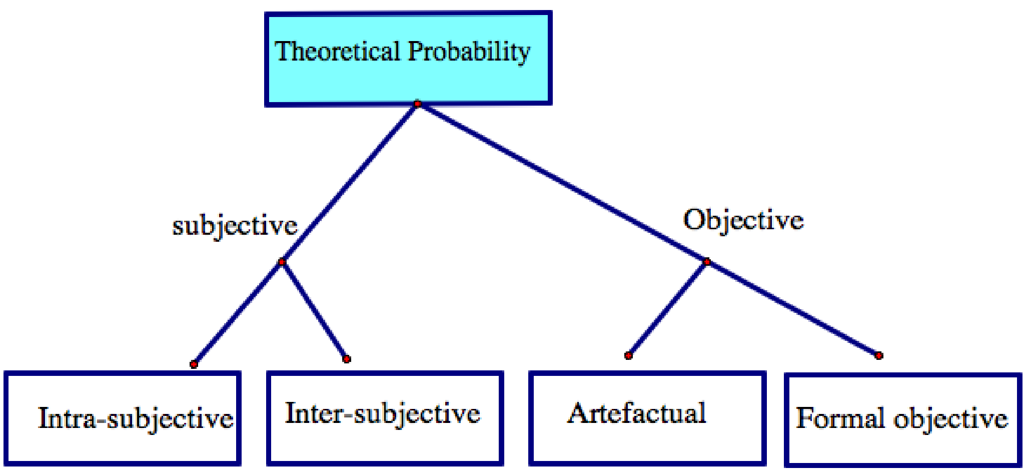
Example of a certain event	Condition
Correct (n=21)	N/A
Incorrect (n=24)	Lack of key information or common sense (n=19) Non-random situation (n=5)

Richness: In richness category I consider the context from which the example is generated. Everyday experience and mathematical experience are the two main contexts that have been looked for. However it is not an exhaustive partitioning of the possible contexts for the examples and also the two are not mutually exclusive since mathematics both comes from (not all of it though) and is applied to the real life. In order to decide on whether the context of an example is mathematical or not, I have looked for evidence of combinatorial reasoning, meaningful use of numbers or standard randomizers such as coin, dice, spinners, urn of balls, etc. 16 examples are marked as including mathematical context and the other 29 examples are describing less mathematical, and more real life situations.

Table 2: richness and objective-subjective duality analysis of the examples (n=45)

Mathematical (n=16)	12 correct 4 incorrect	14 Artefactual Objective 2 Formal Objective 0 Inter-Subjective 0 Intra-subjective
Everyday life (n=29)	10 correct 19 incorrect	3 Artefactual Objective 0 Formal Objective 16 Inter-Subjective 10 Intra-subjective

Subjective-Objective: In the next step, the examples are put into two main categories of objective and subjective inside which four refined categories of "formal objective", "artefactual", "inter-subjective", and "intra-subjective" are recognized.



These expressions are adopted from Gillies (2000) and used as informative and distinctive probability terminologies by Chernoff (2008); here is a very brief description of each:

Artefactuale: "probabilities can be considered as existing in the material world and so as being objective, but they are the result of interaction between humans and nature. Probabilities in coin tossing and other games of chance, as well as the probabilities associated with repeatable experiments in science, are artefactual", Gillies (p.171).

From the 45 examples provided by the participants, 17 examples were identified as referring to artefactual type of probability. Evidence of combinatorial reasoning (e.g. "10 red apples and 1 green apple in a basket, you are guaranteed with picking 1 red apples with 2 chances without replacement"), references to the games of chance (e.g. "Probability of getting heads or tails when tossing a coin is 100%"), and addressing statistical or scientific findings (e.g. "An earthquake here in BC in the next hundred years occurs with 100% chance. The experts have been predicting it for decades but no one knows when it will happen") are used as the main criteria for this group of examples.

Formal objective: Those events that are independent of humans—to the greatest extent that they can be, for example events related to the hypothetical problem of dropping a needle on number line and finding the probability of certain set of numbers being hit (rational numbers for instance) is divorced enough from the human context to be categorized with "formal objective". Within the collected data there were only two incidents of formal objective type of probability:
1) "There is a 100% chance that in flip of a coin it is 50% probability that it flip heads".
2) "The probability of rolling a 1 on a 6-sided die being 1/6".
It could be argued that these examples are more of an artefactual type since they refer to the well-known facts. It is apt to distinguish between the mathematical facts and statistical facts and the mode of inquiry of those. I contend that if the participants consider "Probability of heads or tails are each 50%" as a result of experiment or as what statistics suggests, then these examples are merely artifacts and hence pertain to artefactual probability. But if we look at these facts as results of mathematical theorems (e.g. ($\lim_{n\to\infty}(p(\left|\frac{x}{n}-\frac{1}{2}\right|<\varepsilon))=1$ for the coin tossing experiment), then the two examples above are assigning an objective probability to an event which is far away enough from the human context to sit with Formal objective probability. This itself is a fascinating example of how the perspective of the person who is examining these examples (that would be me in this case) can affect on the probability stance of a single probability assessment.

Inter-subjective: probabilities that represent the degree of belief of a social group that has reached a consensus. In other words it includes probabilities that are assigned on a subjective basis but in the light of some evidence that are clear to a group of people. For example the probability of Sara taking an umbrella on a cloudy November day of Vancouver could very well be assigned on an inter-subjective way. The followings are two examples from the 16 examples identified as bearing indications of Inter-subjective probability. In these examples the probability proposed is perceived (by me) as containing no formal calculation, but close to what might be akin to the belief and knowledge of a group of people.
"There is a 100% chance of having three girls in a lecture room that contains 100 students". This example is marked as incorrect because of the lack of accounting for a possible scenario, which is "an all boy class", but nevertheless it tacitly referrers to the experience of students from their large classes in a typical university/college.
"I am 100% sure that most of the class is right-handed." Once again this example is not referring to any statistical finding or a ratio of right-handed to the total, but reasonably it is believed that most of people are right handed and a class is an appropriate representative of the whole population with regard to this feature.

Intra-subjective: it is more of a personal belief-type of probability. The probability that I'll take the bus tomorrow (and no further evidences or information provided) is put into this category.
Examples include: "There is a 100% probability that I will take the bus home", "Going to bed tonight", and "It is 100% probable that at least 2 people in this class will be born in the same birth month".

It is necessary to note that the four above-mentioned types of probabilities (formal objective, artefactual, inter-subjective, and intra-subjective) mostly describe the extremes and indicate the upper and lower bounds of the probability continuum. In total, 10 examples fit this category all of which are marked as "incorrect" in the first run of examining the examples. This brings us back to the issue of introducing subjective probabilities into the k-12 mathematics curriculum.

When it comes to marking tests or using other forms of evaluation, we need to decide whether it is possible to develop a consistent criteria to mark the students' intra-subjective arguments or not. Or it could be the case that any intra-subjective probability assignment by definition carries a connotation of "wrong or insufficient explanation.

Concluding remarks:

The absence of 'expert' example space (as described in Zazkis & Leikin, 2008) that displays rich variety of expert knowledge is apparent. All of the mathematical examples obtained from the participants are textbook examples of sure events typically used by the teachers at the very first phases of the introduction of the notion. Examples related to basic coin tossing or die rolling events as well as statements such as "the sun will rise tomorrow" (which is yet another textbook example of certain events, p.474 Van de Walle). The later performance of the participants of this study on the probability test (not reflected in this paper) shows that they have a reasonably good grasp of laws of probability and that they have an above average performance with the probability related tasks and problems. Yet their examples of certain events don't reflect the same level of development and expertise. An overwhelming 26 out of 45 example referred to in this study proved to be subjective probability statements, pointing out the fact that the frequency and classical (objective) approaches to probability are less widely applicable than the belief interpretation. A person can hold beliefs about any event, but the frequency interpretation applies only when a well-defined experiment can be repeated and the ratio always converges to the same number. Many events for which we would like to have probabilities clearly do not have probabilities in the frequency and classic sense. For example consider the most frequently mentioned sure event: several participants presented the "I will die" example as an event with 100% probability of happening. Let's try to assign a classical probability to this event: we first need to define a sample space consisting of *equiprobable* events, count the number of events in which "I will die" and divide it by the total number of the events in the sample space. The inherent difficulty in doing so may lie in the idea that the sample space is either S={I will die, I will not die} or S={I will die}. The former is suffering from the absence of equiprobability and the latter is acceptable only if we have made up our mind (in an a-priori fashion) that nothing else is possible and thus the "I will die" event is the 100% sure event. This conceptual difficulty is not specific to the extreme probabilities, subjective aspects of making decisions about assigning or calculating mathematical probability remains the same all over the probability continuum, but they are more noticeable in the case of impossible and certain events.

I propose that students of probability at all levels need to experiment with probability tasks in which they are not only asked to calculate/assign the probability of an event but also they are encouraged to uncover and discuss the underlying assumptions that are made about the event in question and the knowledge of different individuals about the event. The dynamic process of taking in new information and adjusting the previously formed beliefs and judgments creates not only a bridge between frequency based and subjective probability measurements but also creates valuable opportunity for students to develop a new perspective on uncertainty.

References:

Antonini, S., & Presmeg, N., & Mariotti, M., & Zaslavsky, O. (2011). On examples in mathematical thinking and learning. ZDM Mathematics Education. 43, 191-194.

Bills, L., Dreyfus, T., Mason, J., Tsamir, P., Watson, A., & Zaslavsky (2006). Exemplification in mathematics education. In J. Novotna (Ed.), Proceedings of the 30th Conference of the

International Group for the Psychology of Mathematics Education. Prague, Czech Republic: PME

Chernoff, E. J. (2008). The state of probability measurement in mathematics education: A first approximation. Philosophy of Mathematics Education Journal, 23.

Gillies, D. (2000). Philosophical theories of probability. New York: Rutledge.

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realties. In F. K. Lester (Ed.), Second handbook of research on mathematics teaching and learning (pp. 909-955). New York: Macmillan.

Mason, J. (2011). Phenomenology of example construction. ZDM Mathematics Education. 43, 195–204.

Van de Walle, H., & Folk, S., & Karp, K., & Bay-Williams, J. (2011). Elementary and Middle School Mathematics. Pearson. Toronto.

Vinner, S. (2011). The role of examples in the learning of mathematics and in everyday thought processes. ZDM Mathematics Education. 43, 247–256.

Watson, A., & Mason, J. (2005). Mathematics as a constructive activity. Lawrence Erlbaum Associates.

Watson, A., & Chick, H. (2011). Qualities of examples in learning and teaching. ZDM Mathematics Education. 43, 283–294.

Zazkis, R., & Chernoff, E. (2008). What makes a counterexample exemplary? Educational Studies in Mathematics. 68:195-208.

Zazkis, R., & Leikin, R. (2008). Exemplifying definitions: a case of a square. Educ Stud Math. 69, 131–148.

MODEL-OF TO MODEL-FOR IN THE CONTEXT OF RIEMANN SUM

Kritika Chhetri
University of Northern Colorado

Jason Martin
University of Central Arkansas

This research focuses on the cognitive challenges that students face and how they resolve these challenges while transitioning from intuitive reasoning to constructing a more formal mathematical structure of Riemann sum while modeling "real life" contexts. A pair of Calculus I students who had just received instruction on definite integral defined using Riemann sums and illustrated as area participated in ten interviews. They were given three contextual problems related to Riemann sums but were not informed of this relationship. The intent was to observe students' transitioning from "model-of" to "model-for" reasoning based on Gravemeijer and Stephan (2010). Findings indicate that it was not the end results but records of their ways of acting and reasoning about their contextual problem through multiple representations along with real life intervention that served as tools for supporting their transition from "model-of" informal activities to "model-for" more formal mathematical reasoning.

Keywords: Emergent modeling, Riemann sum, Realistic Mathematics Education, Design experiment, Model-of/Model-for

Introduction and Research Questions

Orton (1983) observed that many students who could fluently compute the definite integral could not explain why they needed the definite integral. Sealey (2008) highlighted that an understanding of the structure of the Riemann sum provides a foundation upon which students can understand why definite integrals model various situations found within physics and engineering. Von Korff and Rebello (2012) emphasized that in addition to knowing the appropriate structural elements of the Riemann sum, students also need to have a good understanding of the physical context of their problem situation. For that reason they stressed encouraging students to manifest the structural elements of Riemann sum into multiple representations. Sealey and Oehrtman (2007) also detailed the importance of conceiving of appropriate structural elements of the Riemann sum within contexts in order to complete approximation tasks. Research detailing how students might shift from informal activities to a more formal understanding of the definite integral has leveraged quantitative reasoning and how that reasoning can support a more conceptually accessible formation of the Fundamental Theorem of Calculus (Thompson & Silverman, 2008). But when students come to understand Riemann sums as a model of a particular situation, how does their reasoning about that model influence their reasoning in constructing Riemann sum models of subsequent situations? This study attempts to address how students shift from informal to more formal mathematical reasoning about Riemann sums and definite integrals using an emergent modeling approach. Studies that have detailed the cognitive challenges that students face in modeling contextual problems using Riemann sums (Sealey, 2008) have not explicitly incorporated an emergent modeling approach. This research attempts to answer the following questions. (1) What role do representations play in identifying and resolving challenges as students construct a Riemann sum as a "model of" a contextual approximation problem? (2) How do "records of" and "tools for" aid students while transitioning from "model of" to "model for" reasoning about structural elements of the Riemann sum?

Theoretical Perspective

Emergent modeling is a Realistic Mathematics Education's (Freudenthal, 1973) instructional design heuristic where modeling is viewed as an active organizing process in which models co-evolve as students reorganize intuitive reasoning and construct more formal mathematical reasoning (Gravemeijer, 2002). In emergent modeling, models are more than representations; they are holistic organizing activities including a solution strategy. Sub-models evolve along students' ways of acting and reasoning about their problem situation, and build one after another through an iterative process. Gravemeijer & Stephan (2010) classify the initial informal context-specific reasoning as "model-of", and the more generalized reasoning as "model-for". As students employ their "model-of" reasoning and begin to identify commonalities amongst different situations, their model starts to change character, and slowly takes on a life of its own, with its own identity, and emerges as a new mathematical reality (Gravemeijer & Stephan, 2010). "Real life" refers to situations that are experientially real to students, and mathematical reality implies mathematical reasoning that students access intuitively and experience as their reality (Gravemeijer, 2007; Johnson, 2013). While transitioning from a model-of a context-specific reasoning to a model-for general reasoning, students actively engage in symbolizing their emerging reasoning (Gravemeijer & Stephan, 2010). As students symbolize and revise their reasoning in an iterative process, various signs emerge which take on different roles that can be referred as "record-of", and "tool-for" (Zandieh & Rasmussen, 2010). "Record-of" and "tool-for" are ways of analyzing day-to-day level teaching experiments to emphasize important transitions without the need to associate those transitions to new mathematical realities (Rasmussen & Marrongelle, 2006). While model-of/model-for captures the big picture of how students create a new mathematical reality, record-of and tool-for focuses the symbolizing that students engage in while creating their new mathematical reality. Records-of can emerge as representations (including pictures of situations, symbols, tables of values, graphs, etc.) that embody students' reasoning about a problem situation (Zandieh & Rasmussen, 2010). A tool is something that the student "explicitly recognizes as useful for achieving specific goals" (Rasmussen & Marrongelle, 2006, p. 2). Various records-of and tools-for emerge while transitioning from a model-of/model-for reasoning. In the context of Riemann sums, a record-of students' picture of a dam may later serve as a tool-for subsequent mathematical reasoning about force. Quantitative reasoning (Larson, 2010; Thompson, 2011) serves to support the transition from 'record of' to 'tool for' through initially conceiving of, representing, operating on, coordinating, and identifying new context-specific quantities and later identifying commonalities between quantities in different newfound contexts. When a conceived quantity is specifically attached to an attribute of a problem situation, any representing of this quantity would indicate model-of reasoning, but as one reasons about this quantity within a quantitative structure without referring to a problem situation, that reasoning emerges as a model-for their reasoning about the mathematics.

Methods

Ten interview sessions (50-148 minutes) were conducted with two volunteer first-semester calculus students, Sam and Chris (pseudonyms) after they had been introduced to the definite integral through Riemann sums illustrated as area (Stewart, 2008). Students were given three approximation tasks related to Riemann sums and adapted from Oehrtman's (2008) approximation framework. Of the three tasks, two emphasized finding under and overestimates to total distance traveled based on a table of velocities and a velocity function, respectively

(Figure 1). Tasks 1 and 2 were scaffolded with nine subtasks that included drawing pictures of the actual situation, finding and illustrating error bounds, and graphing. Task 3 entailed approximating total force exerted on a dam with scaffolding mostly removed. After completing all tasks, students were asked to compare and contrast their solution methods in a generalizing activity. Sessions were videotaped to analyze how students modeled their problem situations.

Models were identified based on students' reasoning as exemplified by their representations and verbal utterances. When students directly related their reasoning to a problem situation, this was viewed as model-of reasoning. When students did not tie their reasoning to a particular task, but rather generalized it and then extended it to other situations, this was considered as indicators for "model-for" reasoning. As students transitioned from model-of to model-for, initial representations and corresponding descriptions were identified as records-of their reasoning. Prior records-of reasoning and symbolizing when applied were viewed as indicators of tools-for reasoning. Students' conceptions related to distance/rate/time relationship (DRT) and force/pressure/area relationship (FPA) will be referred to as micro-level models. Conceptions of DRT and FPA, which approximate total distance traveled or total amount of force will be referred to as macro-level models. The micro-level sub-models relate to conceiving of quantities and constructing a multiplicative relationship between those quantities. Macro-level models relate to the idea of adding individual products that have been constructed at the micro-level to approximate a total, total distance traveled or total force exerted on a dam.

Task 1: The table drawn below shows the velocity of a car travelling from Conway to Little Rock. In this activity you will approximate the distance travelled by the car during the first 10 seconds of the car entering the southbound I-40 ramp.

T(s)	0	2	4	6	8	10
V(ft/s)	0	21	34	44	51	56

Task 2: NASA's Q36 Robotic Lunar Rover can travel up to 3 hours on a single charge and has a range of 1.6 miles. After t hours of traveling, its speed is miles per hour given by the function $v(t) = sin\sqrt{9 - t^2}$. In this activity you will approximate the distance travelled by the Lunar Rover in the first two hours.

Task 3: A uniform pressure P applied across a surface area A creates a total force of $F=PA$. The density of water is 1000 kg per cubic meter, so that under water the pressure varies according to depth, d, as $P=1000dg$. In this activity you will approximate the total force of the water exerted on a dam 62 meters wide and extending 25 meters under water.

Figure 1. The three approximation tasks without scaffolding

Results

The results reported here will focus on how students' records-of reasoning about DRT relationships and approximations to total distance in Tasks 1 and 2 emerged as model-for reasoning. Descriptions for distinct DRT and FPA conceptions, together with descriptions of their emerging conceptions for finding approximations to total distance traveled and total force can be found in Tables 1-3. Initially, at the micro-level, Sam and Chris quickly identified varying velocities and the finite amount of data as causing problems with completing Task 1. They said, "Distance changes as time changes" (DRT 1) and a record-of this reasoning is found in Figure 2. After the facilitator prompted them to be "picky" about their picture, they noted amounts of change in distance should vary but they did not attend to the detail of *how* amounts of change in distance would vary. They represented this conception pictorially with increasing changes in distance between every two-second snapshot (Figure 3) and formulaically as "$d = \Delta V \cdot \Delta t$" (DRT

2). After prompted to think about a “real life” situation of a car merging onto an interstate, Sam drew a picture of a moving car (Figure 4) indicating increasing distances between snapshots and said, “It’s [car] where it starts, would be the next place and the next place will be little bit further and further and further.” Then they revised their formula to $d = V_p \Delta t$ to indicate increasing distances between snapshots (DRT 3).

Table 1.
Distinct conceptions during Task 1.

Conception	Description of reasoning
DRT 1: Distance changes as time changes	Omitted explicit detail to amounts of change in velocity. Pictorially represented as a vehicle with constant amounts of changes in distance per two-second intervals.
DRT 2: Distance is change in velocity × change in time	Initially supported by their reasoning that amounts of change in distance vary because of changing velocities. Pictorially represented as a vehicle with decreasing amounts of changes in distance per two second intervals which became a model for distance as $d = \Delta V \cdot \Delta t$.
DRT 3: Distance is constant velocity × change in time	Initially only conceived for a vehicle traveling at constant velocities. Only after adjusting their picture to model a vehicle with increasing amounts of changes in distance and after “supposing” their vehicle as traveling at constant velocities was this conception applied to their context. Formulaically represented as $d = V_p \Delta t$.
Total 1: Total distance approximated by adding distances can be underestimates or inconclusive.	Adding up amounts of change in distances approximates total distance. Coordinated with DRT 2 and then DRT 3. With DRT 3 it was initially represented as $\sum_{p=0}^{5} V_P \Delta t$. For Sam, this sum was an underestimate because the sum would increase towards the exact total distance traveled as more data points were added. For Chris, this sum was inconclusive because the data table did not reveal what happened between data points.
Total 2: Total distance approximated by adding using max. and min. velocities.	Coordinated with DRT 3. They conceived of maximum and minimum velocities over a time-interval as approximations to varying velocity over that interval. Underestimates and overestimates were represented by $\sum_{p=0}^{4} V_P \Delta t$ and $\sum_{p=1}^{5} V_P \Delta t$, respectively.

After establishing DRT 3, at the macro-level, they totaled individual distances to 412 ft, but struggled to identify it as an overestimate to total distance traveled by the car during the first ten seconds. They did not attend to a notion of max. and min. velocity over an interval, which hindered them from reasoning about their approximation as an overestimate. Then the facilitator engaged them in reasoning about 21 ft/s as a max. and a min. velocity over the 0-2 and 2-4 second intervals. However, record of their reasoning about 21 ft/s as a max. or min. velocity over the 0-2 or 2-4 second intervals, respectively, did not immediately support them in justifying why 412 ft was an overestimate to total distance traveled over the 10-second interval. It was only after they had to find the error bound that they exploited their reasoning about 21 ft/s as both a min. and max. velocity over 2-4 and 0-2 second intervals to justify 412 ft and 300 ft as an over and underestimate, respectively. As they proceeded to illustrate distance on a graph of velocity as a function of time, students’ records-of their picture (Figure 4) served as tools for attributing distance between two points as straight-line distances on secant lines to the curve instead of area under the curve. Even though Sam acknowledged that these distances, symbolized as $\sqrt{\Delta t^2 + \Delta V^2}$ on their graph, did not yield correct units, he still grappled to illustrate distance. Even their prior experiences with illustrating Riemann sums from their calculus class did not serve as tools to attend to rectangles of the curve. Only after the facilitator intervened and drew a rectangle over the first interval did Sam attend to distance as area of rectangles and employ his prior reasoning about max. and min. velocity over an interval to categorize rectangles below and above the curve as under and overestimates to distance traveled.

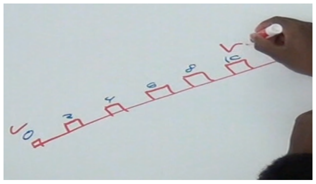


Figure 2. DRT 1

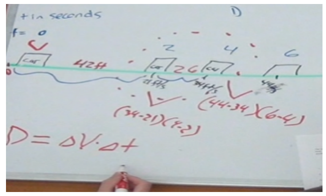


Figure 3. DRT 2

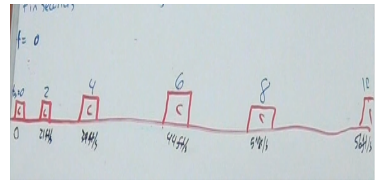


Figure 4. DRT 3

Their ways of reasoning about DRT and Total concepts in Task 1 supported them to engage in similar types of reasoning in Task 2. While they completed Task 1 in about six hours, they completed Task 2 in about three hours (including a new subtask in Task 2 of finding an approximation to a predetermined error bound). Immediately after starting Task 2, Sam asked Chris, "Well, you think the picture looks the same as last time?" and Chris responded, "Yeah. We need to know if the velocity is increasing or decreasing." Sam drew a picture with snapshots of the rover at every half-hour interval (Figure 5), and Chris constructed a table similar to Task 1 using the rover's velocity function. Records of reasoning about picture from Task 1 became tools for Sam, while records of reasoning about a table from Task 1 became tools for Chris.

Figure 5. Sam's first picture of a Lunar Rover traveling for two hours.

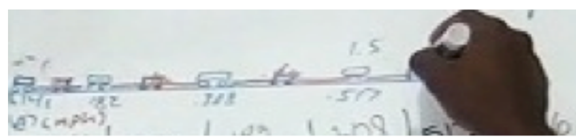


Figure 6. Chris constructing more snapshots of Rover.

After being asked, "Where is distance?" they noted distances between snapshots on their picture and proclaimed that finding total distance was, "the same as what we did last time" and wrote total distance as $\sum V_{(t)}\Delta t$. While computing approximations, they readily attended to notions of max. and min. velocities to support their reasoning for confirming their approximations as under and overestimates to total distance traveled by the rover in two hours. They graphed distance as areas of rectangles above and below the curve. They demonstrated the number of snapshots as going to infinity by drawing more snapshots (Figure 6), and formulaically as "$D = \lim_{n\to\infty} \sum V_t(\frac{t_f - t_i}{n})$." Task 2 had a new subtask where they had to find an error bound to within one foot. Although they reasoned that more snapshots would yield a smaller error bound, they struggled to find an *n* that would give the desired accuracy. Later Chris wrote "$V_n(\frac{t_f-t_i}{n})$ - $V_1(\frac{t_f-t_i}{n})$ = 0.00189 miles" and reasoned that they could solve for *n* by subtracting, "The last value here [pointing to total overestimate] minus the first value here [pointing to total underestimate] actually because others [values] are going to cancel." Sam agreed with Chris and found *n* to be 6832 (later rounded to 6833) snapshots. They wrote, "As our error bound decreases we approximate closer to the actual distance. In this case, we started with five snapshots, and we had an error bound of 0.323 miles. When we increased the number of snapshots to 6832, we had an error bound of 0.000189 miles."

Both students referred to their prior reasoning about DRT and Total conceptions to reason about pressure and total force in Task 3 but each had their own sets of challenges and resolutions. Therefore, Chris and Sam's FPA and Total conceptions have been classified into Tables 2 and 3, respectively. Although Chris readily attended to changing pressure based on the provided formula, he struggled to understand the relationship between force, pressure and depth, which hindered him in moving ahead. After Sam explained to him about his experience with swimming, Chris was able to make connection between pressure and depth. Then records of his reasoning about max. and min. velocity served as tools to construct and reason about total under and overestimate force exerted on the dam. For example, he said the following:

> We calculate the force […] different forces and […] summation of those the max. forces and the min. forces, the difference between the max. forces and the min. forces [...] actual force will be in between them. And we can make it the exact force if we apply the limits.

Unlike Chris, Sam readily attributed the relationship between pressure and depth to reason why total force was not just the product of force and area. He employed his real life swimming experience to reason that pressure would increase with depth, and said, "Pressure here cannot be multiplied by only an area because depth must be taken into consideration." Although he intuited that different depths would result different pressures, he struggled to formulate his reasoning about total force. Then Chris helped him in symbolizing total under and overestimates to force and also in solving for an *n* to within 5000 N error bound.

Table 2.
Chris' FPA and Total Force Conceptions.

Chris' FPA Conception	Description of reasoning
FPA 1: Pressure is changing	Formula served as tools for interpreting pressure to be changing.
FPA 2: Pressure depends on depth	Initially modeled by the provided formula and then reinforced by Sam's reasoning about his "real life" swimming experience. Pictorially represented by increasing pressure along the depth of a dam.
FPA 3: Pressure at a depth	Reasoning about pressure from his table supplemented with Sam's over and underestimate expression, Chris conceived of pressure at a point and verbalized his notion of minimum and maximum pressure at a point [depth].
FPA 4: Force= Pressure at a point [depth] * Area	Reasoning about min and max pressure from their picture and table allowed Chris to reason about force as a product of pressure at a point and its area.
Total 1: Total force approximated by adding using max and min pressures at different depths.	Coordinated with FPA 4, and facilitated with their prior reasoning of minimum and maximum velocity, they conceived of maximum and minimum pressure at a depth as approximations to varying pressures along the depth. Underestimates and overestimates were represented by $\sum_{b=0}^{4} P_{b*}62 * \Delta d$ and $\sum_{b=1}^{5} P_b 62 * \Delta d$, respectively.

In their generalizing activity, they highlighted that their ways of reasoning was similar in all of their tasks. Chris said, "If it was the force […] we considered a lot of number[s] of pressure to make our error smaller and to find the exact, we applied the number of those points tend to infinity…[points to Task 1] for the distance we considered the number of snapshots tends to infinity." To represent their approximation in each of their contexts, Sam spontaneously drew rectangles (Figure 7) and labeled them as their approximations for their tasks. He said, "Each time we found the area […] our approximation was the area between two other variables […] Task 1 would be velocity times time equals our displacement … and then the other would be pressure times area equals force." He pointed to rectangles of prior graphs and emphasized that summation of those rectangles would give him the approximation. They symbolized their approximation as $\sum_{p=0}^{n} a_p \left(\frac{b_{f-}b_i}{n}\right)$. When asked if they saw anything common between their contexts, Sam immediately referred to the multiplicative structure of two variables as being common in all of the tasks and said, "Anything that's just one thing times another is gonna have that [points to rectangles] area and you can [...] get the exact amount of it by taking an infinite number of snapshots of it" and drew Figures 8 and 9 to illustrate the multiplicative relationship for two new contexts he provided, voltage and force. Finally, they were asked if they had seen something similar in their calculus class, Chris replied, "This would be the integral, integral of the products of two things." After writing definite integrals for each of their tasks, Sam expressed that graphically the definite integrals represent the exact areas under the curves.

Table 3.
Sam's FPA and Total Force Conceptions.

Sam's FPA Conception	Description of reasoning
FPA 1: Pressure increases with depth	Aided with algebraic expression, $F=\rho g d A$ and supplemented with his "real life" swimming experience, Sam reasoned that pressure increases with depth and modeled his reasoning through his picture.
Total 1: Total force is the sum of $\rho g \Delta d A$	Reasoning with and about table and picture served as tools for reasoning about total force as the sum of $\rho g d A$.
Total 2: Total force is the sum of $\rho g n \Delta d A$	Organized his Total 1 to model a dynamic pressure and reasoned that multiplying with n would provide him with a constantly changing pressure.
Total 3: Total force approximated by adding using max and min pressures at different depths.	Organized his reasoning about $n\Delta d$ with Chris' pressure at a point and constructed his underestimate and overestimate force as $\sum_{b=0}^{4} P_{b*}62 * \Delta d$ and $\sum_{b=1}^{5} P_b 62 * \Delta d$, respectively.

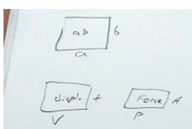


Figure 7. Approximations as rectangles.

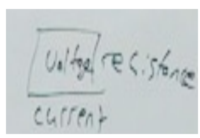


Figure 8. Voltage as a product of resistance and current.

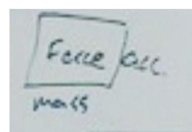


Figure 9. Force as a product of mass and acceleration.

Discussion

We note that it was not merely the end results, but elements of students' reasoning, including their solution strategy that served as tools for reasoning more effectively about subsequent tasks. Being able to imagine a constant quantity as approximating a varying quantity was crucial for them to reason about their approximations. Manifesting problems and solutions into multiple representations supported them in making sense of their context and in building rich connections amongst their representations. Records of their picture, graph, table, and formulaic expressions from Task 1 served as reference points for them to make connections between their two tasks as they conceived of, represented, and related relevant quantities. Real life intervention enabled students to understand their context and facilitated them to organize their informal activities into more formal mathematical reasoning of Riemann sums. Since students spontaneously extended their micro-level conception of product structure to novel situations (e.g. Figure 8, Figure 9), the results suggest that they conceived of their product structure as a new mathematical reality. With just three tasks, it is hard to claim if or exactly when students transitioned from model-of to model-for reasoning at the macro-level since they did not expand their reasoning of Riemann sums as an *informal* activity for some other task. Because no such extension of Riemann sum was observed, it does not mean that it had not become a model-for, just that this type of reasoning was not clearly evidenced. Using additional contexts to clearly observe if students conceived of the Riemann sum as their new reality could extend our knowledge of if their reasoning about Riemann sum had emerged as model-of reasoning. These observations guide plans for future research designed to support students in an observable transition from model-of/model-for reasoning at the macro-level. We acknowledge that work with one pair of students may not necessarily generalize to others, but even so, this study highlights that leveraging real life experiences and utilizing multiple representations can support students in a deeper understanding of Riemann sum and in realizing a need for definite integral. While Sealey (2008) detailed students' understanding of Riemann sums as layers and Von Korff and Rebello (2011) demonstrated how multiple representations and activities rooted in physical contexts can support reasoning, this study provides more detail into how *coming to* understand physical contexts through evolving models of a particular situation, even "incorrect" models, can subsequently support the generalization of elements of a learning process about Riemann sums.

References

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht, Holland: D. Reidel

Gravemeijer, K. (2002). Emergent modeling as the basis for an instructional sequence on data analysis. In *Proceedings of the 6th International Conference on Teaching Statistics*. Voorburg, the Netherlands.

Gravemeijer, K., & Stephan, M. (2010). Emergent models as an instructional design heuristic. In Gravemeijer et al. (Ed.), *Symbolizing, modeling, and instruction design* (pp. 145-169). Mahwah, NJ: Erlbaum and Associates.

Gravemeijer, K. (2007). Emergent modelling as a precursor to mathematical modeling. In W. Blum, P.L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education. The 14th ICMI Study* (pp. 137-144). New York: Springer.

Johnson, E. (2013). Implication of Realistic Mathematics Education for analyzing student learning. *In Proceedings of the 16th Conference on Research in Undergraduate Mathematics Education*. Denver, CO: University of Northern Colorado.

Larson, C. (2010). Modeling and quantitative reasoning: The summer jobs problem. In R. Lesh, C. R. Haines, P. L.Galbraith & A. Hurford (Eds.), *Modeling Students' Mathematical Modeling Competencies: ICTMA 13* (pp. 111-118). New York, NY: Springer.

Oehrtman, M. (2008). Layers of Abstraction: Theory and design for the instruction of limit concepts. In M. P. Carlson, C. Rasmussen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* (Vol. 73, pp. 65-80). Washington, DC: Mathematical Association of America.

Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics, 14*(1), 1-18.

Rasmussen, C. & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics into instruction. *Journal for Research in Mathematics Education*, 37, 388-420.

Sealey, V. L. (2008). *Calculus students' assimilation of the Riemann integral into a previously established limit structure.* Dissertation, Arizona State University, United States, Arizona.

Sealey, V., & Oehrtman, M. (2007). Calculus students' assimilation of the Riemann integral. *In Proceedings of the 10th Conference on Research in Undergraduate Mathematics Education*. San Diego, CA: San Diego State University.

Stewart, J. (2008). *Calculus: Early Transcendentals* (6th ed.). Belmont, CA: Thomson Learning, Inc.

Thompson, P. (1994). Images of rate and operational understanding of the Fundamental Theorem of Calculus. *Educational Studies in Mathematics*, 26(2-3), 229-274.

Thompson, P. (2011). Quantitative reasoning and mathematical modeling. In L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspective and directions for collaborative research in mathematics education* (pp. 33-57). Laramie, WY: University of Wyoming Press.

Thompson, P., & Silverman, J. (2008). The concept of accumulation in calculus. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (Vol. 73, pp. 117-131). Washington, DC: Mathematical Association of America.

Von Korff, J., & Rebello, N. S. (2012). Teaching integration with layers and representations: A case study. *Physical Review Special Topics - Physics Education Research, 8*(1), 010125.

Zandieh, M., & Rasmussen, C. (2010). Defining as a mathematical activity: A framework for characterizing progress from informal to more formal ways of reasoning. *Journal of Mathematical Behavior, 29*, 57-75

43

USING A FRAMING AND RESOURCES FRAMEWORK FOR ANALYZING STUDENT THINKING ABOUT MATRIX MULTIPLICATION

Warren Christensen
North Dakota State University

A student who has completed both Linear Algebra and Quantum Mechanics should have a wealth of conceptual and procedural knowledge that has been obtained from mathematics and physics classes. However in practice, students seem to struggle with this task. This investigation casts light on students' thinking about matrix multiplication and how their thinking appears to be influenced by their framing of the problem as either a mathematics or physics question. Using Framing and Resources as a theoretical lens can provide insight into the ideas and concepts that a student accesses from domains of mathematics and physics. Using lexicon analysis, it appears the student shifts from a "mathematical frame" to a "physics frame" and back again, but struggles to successfully transfer concepts between these two frames. I will highlight the markers for these frame shifts and demonstrate why framing and resources is the appropriate lens for this investigation.

Key words: Linear Algebra, Physics, Quantum Mechanics, Interdisciplinary, Interviews

Over the past decades, a great deal of research has taken place within both Mathematics departments and Physics departments on the learning and teaching of these respective fields. A growing number of researchers are interested in tapping the knowledge base of both the communities of Physics Education Research (PER) and Research in Undergraduate Mathematics Education (RUME). The present study builds on the work of Henderson et al. (2010), which used a theoretical framework of symbol sense (Arcavi 1994) to investigate students' reasoning on concepts of Linear Algebra. The current study uses framing and resources as a theoretical framework to investigate students thinking about matrix multiplication after a course in Quantum Mechanics (Hammer, Elby, Scherr & Redish, 2005).

Resources and framing are fairly novel approaches to research for the RUME community, but are prevalent in PER. Within this theoretical framework, individuals have different types of resources. For example, *conceptual resources* deal with understanding physical phenomena, such as one's understanding of the concept of force and Newton's Laws. The resources that the student activates in any particular situation depend on how they *frame* the problem they are considering, that is, how they answer the question, "What is going on here?" (Hammer et al. 2005). Frames are locally coherent sets of resource activations. The process of learning involves forming these sets of activations, and then, once formed, using these frames in settings where they seem applicable. This study uses lexicon analysis to make claims about how a student is framing a particular problem, and what resources that student is activating during a one-on-one interview. Specifically, we identify episodes during the interview when the student has a high density of discipline-specific terms (either mathematics or physics) and then investigate what the student is describing, what triggered the cascade of language, and the appropriate and correctness of their statements.

The case study that will be presented demonstrates the power of this theoretical framework. A student can be seen to answer questions correctly and incorrectly, in what would appear to be two different frames and is unable to use the correct ideas from each frame together to resolve inconsistencies and troubling issues. Discussion will focus on these theoretical frameworks as a tool for better understanding student thinking.

References

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the learning of mathematics, 14*(3), 24-35.

Hammer, D., Elby, A., Scherr, R. E., and Redish, E. F. (2005) Resources, Framing and Transfer, in *Transfer Of Learning: Research And Perspectives (Current Perspectives on Cognition, Learning, and Instruction)*. Information Age Publishing. Ed. Jose Mestre.

Henderson, F., Rasmussen, C., Zandieh, M. Wawro, M. and Sweeney, G. (2010). Symbol sense in linear algebra. Paper presented at the Thirteenth Conference on Research in Undergraduate Mathematics Education, Raleigh, NC.

Undergraduate Students' Stochastic Understanding of Probability Distribution

Darcy Conant
Notre Dame of Maryland University

Stochastic conceptions undergird development of conceptual connections between probability and statistics and support development of a principled understanding (Greeno, 1978) of probability distribution. This study employed mixed research methods to investigate the impact of an instructional course intervention designed to support development of stochastic understanding of probability distribution. Instructional supports consisted of supplemental lab assignments comprised of anticipatory tasks designed to engage students in coordinating thinking about complementary probabilistic and statistical notions along a hypothetical learning trajectory aimed at development of stochastic understanding of probability distribution. Participants were 184 undergraduate students enrolled in a lecture/recitation, calculus-based, introductory probability and statistics course. Results of quantitative analyses showed completion of stochastic lab assignments had a statistically significant impact on students' stochastic understanding of probability distribution. Student interviews revealed those who held stochastic conceptions also indicted integrated reasoning related to probability, variability, and distribution and presented images supporting principled understanding of probability distribution.

Key Words: Probability Distribution, Stochastic Reasoning, Understanding Probability, Simulations.

Introduction and Theoretical Background

Many college-level students struggle with probabilistic and statistical reasoning (Artigue, Batanero, & Kent, 2007; Jones, Langrall, & Mooney, 2007; Shaughnessy, 1992, 2007). Research shows that after a first course in probability and statistics most students do not understand the reasoning involved in making a statistical inference nor the reasoning required for interpretation of the results (Batanero, Tauber, & Sanchez, 2004; Reaburn, 2011; Smith, 2008). Liu and Thompson (2007) found that stochastic conceptions of probability support an understanding of statistical inference. A stochastic conception of probability involves understanding the stochastic nature of random phenomena, understanding how probability is used to model random phenomena, and understanding how probability models are used to make formal statistical inferences (Steinbring, 1991). The purpose of this study was to investigate the impact of an instructional intervention designed to support development of stochastic reasoning by addressing the following research question: What is the impact of an instructional intervention designed to support the development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory, calculus-based, probability and statistics course?

Among prior research studies addressing students' understanding of probability, only a few investigated students' stochastic conceptions of probability and only a few involved participants who had earned college-level credits for calculus. Research shows that post-calculus students, who were either currently enrolled in or had recently completed an introductory probability and statistics course, demonstrated evidence of probabilistic thinking that was aligned with novice thinking evidenced by high school students and college-level students in algebra-based introductory probability and statistics classes (Abrahamson, 2007; Abrahamson & Wilensky, 2007; Barragues, Guisasola, & Morais, 2007; Lunsford, Rowell, & Goodson-Espy, 2006). Research investigating undergraduate post-calculus students' conceptions of probability distribution found that learners exhibited difficulty understanding probability models and struggled to discriminate between empirical distributions and theoretical distributions (Abrahamson, 2007; Batanero, Godino, & Roa, 2004; Lunsford, et al, 2006). Prior research suggests that after completing an introductory, calculus-based, probability and statistics course, most students were comfortable with mathematical procedures and had mastered algorithmic techniques, but lacked stochastic conceptions and a deep conceptual understanding of probability distribution.

Although not conducted in a classroom, the work of Abrahamson (2007) indicates that post-calculus learners can consolidate their intuitive notions of probability with their formal mathematical knowledge in the context of probability distribution. Abrahamson (2007) found that individuals were able to coordinate their thinking about relationships between empirical distributions and theoretical distributions as a result of engaging with interactive models in a computer environment. Other research points to the promise of learners' engagement in simulation tasks utilizing a computer-based, dynamic statistical environment as a means towards facilitating development of notions of sampling distribution, variability, and inferential reasoning (Meletiou-Mavrotheris, 2003; Sanchez & Inzunsa, 2006).

Theoretical Framework

Drawing on constructivist and situated perspectives of learning, the author designed a model which frames development of stochastic understanding probability distribution (Figure 1). This model focuses on individual understanding that is built through learning experiences, which are impacted by the learner, the teacher(s), the instructional material, and peers. Notions of stochastic reasoning are built on experiences with stochastic processes. Stochastic reasoning is crucial for development of understandings which ground aspects of probabilistic and statistical thinking that are essential to understanding probability distribution.

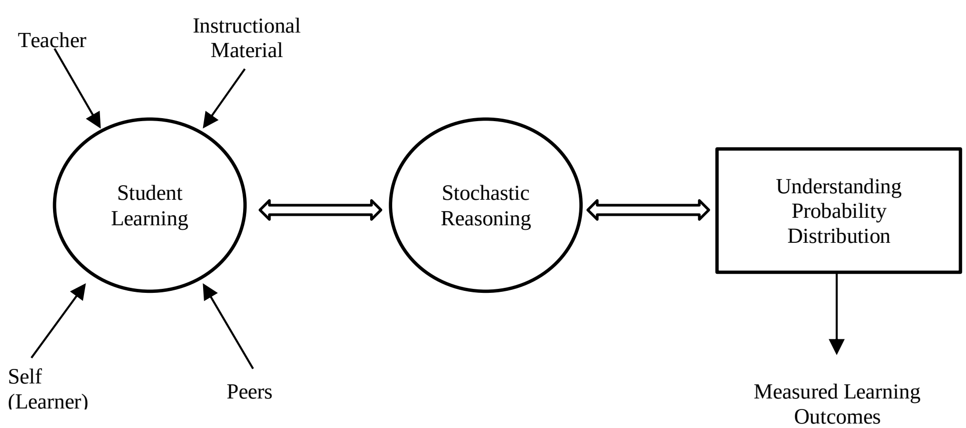


Figure 1. Model of theoretical framework for development of stochastic understanding of probability distribution.

Three overarching constructs frame an understanding of probability distribution: probability, variability, and distribution (Table 1). Understanding probability distribution means understanding connections between the constructs of probability, variability, and distribution, as well as understanding connections among notions within each construct and across the constructs (Kapadia & Borovcnik, 1991). This means that understanding variability is not exclusive of understanding probability or understanding distribution. A deep understanding involves understandings associated with the notions found within the constructs of probability, variability, and distribution, and this deep understanding is fostered by connected conceptions of probability, variability, and distribution.

Table 1
Framework for Understanding of Probability Distribution

Probability Distribution		
Probability	Variability	Distribution
Coordination of empirical and theoretical probability	Randomness and random variability	Distribution of random variable
Random variable	Law of large numbers	Parameterization of distribution model
Sample space	Unit-to-unit variability	Distribution of sample versus population distribution
Independence versus dependence	Sampling variability	Sampling distribution
Model for inference	Variability of sample statistics and the central limit theorem	

Methods

The study employed a mix of both quantitative and qualitative research methods to examine students' understandings that resulted from an instructional intervention in a quasi-experimental, treatment-control setting. A sequential exploratory design (Tashakkori & Teddlie, 2003) allowed elaboration and enhancement of quantitative-based findings with the incorporation of qualitative data in the context of an integrated interpretation. This design incorporated five phases: (1) collection of quantitative data; (2) cursory analysis of conceptual assessment for purposes of interviewee selection, (3) collection and analysis of qualitative data; (4) use of the qualitative findings to inform coding and analysis of conceptual assessment; (5) final analysis and interpretation of both quantitative and qualitative data. Quantitative data sources included a student background survey, an assessment of stochastic conceptions, confidence interval items drawn from ARTIST topic scales (Garfield, delMas, & Chance, 2006), and comprehensive final course examinations. Qualitative data consisted of end-of-course interviews with 12 student volunteers.

Students were assigned to either the treatment group or control group via their course registration and enrollment in a discussion section associated with one of two large lecture classes. Students in the treatment group received supplemental lab assignments aimed at the development of stochastic reasoning in the context of probability distribution. The stochastic lab assignments were designed to develop stochastic anticipations (Simon, 2013) and support development of stochastic understanding of probability distribution along a hypothetical learning trajectory (Simon, 1995) adapted from Liu and Thompson (2007). Students in the stochastic-reasoning lab group engaged in activities which utilized *Fathom* (Finzer, 2007) software to run virtual simulations and to represent distribution of outcomes resulting from the simulations. Students in the control group received supplemental lab assignments consisting of a review of calculus content used in this introductory probability and statistics course. In order to measure learning outcomes, all students completed the same stochastic reasoning assessment in the form of a conceptual quiz which consisted of two problems. Each quiz problem involved a different context representing a probabilistic situation that could be approached stochastically or nonstochastically. Problem one, the "hospital problem", has been used in a number of previous studies which investigated undergraduate students' understanding of probability (Fischbein &

Schnarch, 2007; Reaburn, 2011; Tversky & Kahneman, 1974; Watson, 2000). Problem two was adapted from a question Liu and Thompson (2007) posed to secondary mathematics teachers.

Show all your work and explain your reasoning.

1. A town has two hospitals. On the average, there are 45 babies delivered each day in the larger hospital. The smaller hospital has about 15 births each day. Fifty percent of all babies born in the town are boys. In one year each hospital recorded those days in which the number of boys born was 60% or more of the total deliveries for that day in that hospital. Is it more likely that the larger hospital recorded more such days, that the smaller hospital did, or that the two hospitals roughly recorded the same number of such days? Explain your reasoning.

2. Anthony works at a theater, taking tickets for one movie per night at a theater that holds 250 people. The town has 30,000 people. He estimates that he knows 300 of them by name. Anthony noticed that he often saw at least two people he knew. Assume that people are not coming to the theater because they know Anthony and there is nothing special about the type of movie. Is it in fact unusual that at least two people Anthony knows attend the movie? Give an explanation of your reasoning and be sure to address issues of randomness and distribution.

Figure 2. Stochastic Conception Quiz

Results

Qualitative analyses indicated evidence of stochastic reasoning across three hierarchical categories (Liu & Thompson, 2007): image of a repeatable process, image of specification of conditions, and image of distribution. Evidence for stochastic conception presented in Table 2 was gleaned from interview participants' written work on the conceptual quiz, as well as their responses to interview probes.

Table 2
Summary of Interview-Based Evidence for Stochastic Conception

Image for Repeatable Process	
	Indicates understanding that the process is repeated under essentially the same conditions
	Indicates the repeatable process yields outcomes and *describes outcomes of the process*
	Indicates thinking about a repeated experiment
	Indicates that repetition of process results in repeated sampling
	Connects thinking about the process to a model
	May connect thinking about the process to running a simulation
Image for Specification of Conditions	
	Indicates that repetition of the process yields a collection of variable outcomes
	Assumes outcomes are independent
	Describes a sampling process where each selection is equally likely
	Indicates the sampling process produces samples that are representative of population
	Indicates that variability in outcomes is related to sample size
	Indicates conceiving of conditions of the process in relation to an underlying model
Images for Distribution of Outcomes	
	Indicates thinking about a distribution of outcomes
	Connects thinking about expectation to variability
	Attends to variability when thinking about distribution of outcomes; connects notions of distribution variance, shape, and sample size
	Indicates thinking about the law of large numbers in relation to stabilization of frequencies over a large number of repetitions of the process
	Indicates thinking about an underlying distribution model
	Quantifies "unusual" *as deviation from expectation in terms of a distribution modal*

Note. Italicized images were only evidenced as a result of the interview. Images not italicized were evidenced as a result of the interview and via student's individual written answers for the conceptual quiz.

Three interview participants consistently exhibited stochastic reasoning across all three stochastic reasoning categories for both problem contexts and throughout interview probes and were characterized as holding a stochastic conception. For example, Student 113 explained:

> So, the way I thought about it was, you know, if I flipped a coin 45 times, and in another experiment I flip a coin 15 times. Which one will more likely to get over 60% heads? So I was thinking very quickly you'll see that it would be easier through the things we were doing all semester running these experiments through *Fathom*. …You'd see that you could have, if you only ran 15 trials, you might end up with a result that definitely far away from 50-50.
>
> Once your sample gets bigger, larger and larger and larger, it very, very, very seldom that it'll deviate from, you know, from the expected, or what's expected in the population. You know, like we saw in the *Fathom* experiment. You know, you run the trial in *Fathom* thousands and thousands of times and the line is really, really, really close to what you expect to be, 0.5. So, I sorta felt like 45 is like getting closer to law of large numbers than 15 is. So that was another thing I was thinking about.

Three interview participants exhibited stochastic reasoning across all three categories for only one of the two problem contexts and were characterized as holding a situational conception. The remaining interview participants either indicated no images related to a stochastic conception (characterized as no image) or only indicated an image of repeatable process for either one or both problem contexts (characterized as nonstochastic conception). Table 3 summarizes the 12 interview participants' stochastic conceptions evidenced on the conceptual quiz and interviews.

Table 3
Distribution of Student Identification Numbers over Categories of Stochastic Conceptions

	Stochastic	Situational	Nonstochastic	No image
Interview Participants' Student ID numbers	113 214 329	313 325 817	711 715 1016	428 518 811

Results of the qualitative analysis informed development of a rubric used for scoring all study participants' work on the conceptual assessment. Images of stochastic reasoning evidenced on interview students' written conceptual quiz answers were used to define the rubric, which yielded a maximum score of 14 possible points. The overall distribution of all students' scores on the stochastic reasoning conceptual assessment ranged between 0 and 13 with a median score of 2. This distribution was highly skewed (M = 2.74, SD = 3.26). The distribution of scores for each of the treatment and control groups was also skewed. Students in the stochastic reasoning (SR) group had a median stochastic score of 3, while students in the calculus review (CR) group had a median stochastic score of 0.5. The mean stochastic score for the SR group (M = 3.87, SD = 3.718) was significantly higher than the mean stochastic score for the CR group (M = 1.65, SD = 2.290); [$t(143) = 4.36$, $p <.000$].

A binary variable, stochastic image, was defined to indicate whether or not a student presented evidence of stochastic reasoning on their written work for the conceptual assessment. Table 4 shows results of a logistic regression model fit to the data to predict stochastic image and to investigate the relationship between lab group (treatment/control intervention) and the likelihood a student presented evidence of a stochastic conception. The overall logistic regression model was significantly more effective than the null model in predicting stochastic image, $\chi^2(5, n = 145) = 33.88$, $p < .001$. The odds of a student in the stochastic reasoning lab group presenting a stochastic image were 3.80 times greater than the odds for a student in the calculus review lab group (p = .003). The odds of a student enrolled in Lecture A presenting a stochastic image were 5.89 times greater than the odds for a student in Lecture B (p < .001). The variables used as controls for prior achievement and mathematical and statistical background were not significantly related to stochastic image.

Table 4
Summary of Logistic Regression Analysis Predicting Stochastic Image

Variable	*B*	*SE*	*OR*	95% CI	Wald Statistic	*p*
Group	1.341	.452	3.798	[1.565, 9.217]	8.704	.003
PriorMath	.788	.437	2.199	[0.943, 5.179]	3.252	.071
PriorStat	.105	.442	1.111	[0.467, 2.640]	0.056	.812
Lecture	1.773	.463	5.890	[2.377, 14.595]	14.668	<.001
AllLabsCompleted	.354	.534	1.424	[0.501, 4.053]	0.439	.507
Constant	−3.510	.703	0.030		24.913	<.001

The logistic regression model using the same explanatory variables with an added interaction, all labs completed by group, showed the interaction was significantly related to stochastic image (p = .026). An investigation of this interaction revealed that students in the SR group who completed all of the lab assignments had a higher mean stochastic image than students in the CR group who completed all of the lab assignments (Table 5). Furthermore, the mean stochastic image for students in the SR group who did not complete all of the lab assignments was similar to the mean for students in the CR group who completed all of the lab assignments, as well as the mean for students in the CR group who did not complete all of the lab assignments. These results indicate the stochastic reasoning lab assignments promoted students' movement along the hypothetical learning trajectory. Students who completed all of the stochastic reasoning labs showed a higher propensity towards indicating a stochastic understanding of probability distribution.

Table 5
Mean Stochastic Image by Completion of All Labs and Proportion of Treatment Group

	Lab Assignment Group	
	SR	CR
All labs completed		
Mean stochastic image	0.455	0.128
n	55	48
Not all labs completed		
Mean stochastic image	0.188	0.154
n	16	26

Note. Stochastic Image is dichotomous with 1 = Stochastic or Situational and 0 = Nonstochastic or No Image.

Taken together, qualitative and quantitative results showed that the instructional intervention designed to support development of stochastic understanding of probability distribution incited students' stochastic thinking and promoted a stochastic understanding of probability distribution. Students who held a stochastic conception of probability distribution were able to apply their theoretical understanding of probability distribution to differing empirical contexts. These students demonstrated evidence of thinking that included both a modeling perspective of distribution and a data-centric perspective of distribution (Peters, 2011), and they consistently coordinated an experimental perspective of probability with a theoretical perspective of probability. Some students demonstrated situational stochastic conceptions which were context dependent. For these students, stochastic thinking was evident, but their stochastic reasoning was inconsistent across differing problem contexts. Thus, their stochastic reasoning appeared to be more tenuous. When prompted during the interview, students with situational stochastic conceptions coordinated a modeling perspective of distribution with a data-centric perspective of distribution.

Implications

The results of this study show that particular anticipations and perspectives of distribution are related to a stochastic conception of probability distribution. Students who were characterized as holding stochastic conceptions attended to the notions of variability which included both a data-centric perspective and a modeling perspective of variability. The three interview students who were characterized as holding a stochastic conception of probability distribution spontaneously indicated thinking about a model in relation to the problem situations presented on the conceptual assessment. More robust stochastic conceptions include strong conceptual links between empirical and theoretical distributions. The implication is that coordination of thinking about empirical distributions with probability distribution models could be an essential aspect of stochastic thinking and a principled understanding of probability distribution.

The results of this study also imply that conceptions of repeated sampling are important to a stochastic conception of probability distribution. All students who were characterized as holding stochastic conceptions indicated thinking about repeated sampling in relation to the repeatable process in addition to thinking about repeating an experiment. While some of the students who did not hold stochastic conceptions mentioned sampling, it was clear they were not thinking of a sampling process, and none mentioned thinking about sampling in relation to repetition of an experiment. Furthermore, students who did not hold stochastic conceptions did not indicate a perception of variability in outcomes, but indicated thinking that focused on formulas and calculations. The implication is that anticipation of repeating sampling in relation to a repeatable process is a necessary component of stochastic thinking.

Yet another implication is that students need experiences with distributions of data generated via random processes along with experiences that support development of a modeling perspective of distribution. Experiences with distributions of data should include activities which support development of stochastic anticipations of repeated sampling. Furthermore, these experiences should also include supports which undergird connections between repeated sampling and probability experiments. The implication is that these kinds of experiences with simulations in *Fathom* appear to support development of normative conceptions of randomness and random phenomena which undergird stochastic conceptions of probability distribution.

The results of this study show that the type of task and design of the instruction matter and have implications for curriculum and instruction in probability and statistics. An objective of the stochastic-reasoning lab assignments was to elicit development of stochastic reasoning by means of anticipatory tasks (Simon, 2013), which were designed to prepare students to learn by promoting development of stochastic anticipations. The results imply that the design of anticipatory activities along a hypothetical learning trajectory (Simon, 1995) aimed at stochastic understanding of probability distribution can incite students' stochastic conceptions and potentially change students' thinking. Because stochastic conceptions support thinking about statistical inference and modeling variation, the implication is that instruction aimed at development of stochastic conceptions should be included in the college-level curriculum.

References

Abrahamson, D. (2007). *The real world as a trick question: Undergraduate statistics majors' construction-based modeling of probability.* Paper presented at the American Education Research Association, Chicago, IL.

Abrahamson, D., & Wilensky, U. (2007). Learning axes and bridging tools in a technology-based design for statistics. *International Journal of Computers for Mathematical Learning, 12*(1), 23-55.

Artigue, M., Batanero, C., & Kent, P. (2007). Mathematics thinking and learning at the post-secondary level. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age Publishing.

Barragues, J., Guisasola, J., & Morais, A. (2007). Which event is more likely? Estimating probability by university engineering students. *Recherches en Didactique des Mathematiques, 27*(1), 45-76.

Batanero, C., Godino, J., & Roa, R. (2004). Training teachers to teach probability *Journal of Statistics Education, 12*(1).

Batanero, C., Tauber, L., & Sanchez, V. (2004). Students' reasoning about the normal distribution. In D. Ben-Zavi, & Garfield, J. (Ed.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Garfield, J, delMas, R, & Chance, B. (2006). The Assessment Resource Tools for Improving Statistical Thinking (ARTIST) Project. Retrieved from https://apps3.cehd.umn.edu/artist/scales.html

Greeno, J. (1978). Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist*, 12(3), 262-283.

Finzer, W. (2007). Fathom (2.1) [Software]. Emeryville, CA: Key Curriculum Press.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. Journal for Research in Mathematics Education, 28(1), 96-105.

Kapadia, R., & Borovcnik, M. (Eds.). (1991). *Chance encounters: Probability in education*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Jones, G., Langrall, C., & Mooney, E. (2007). Research in probability. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age Publishing.

Liu, Y., & Thompson, P. (2007). Teachers' Understandings of Probability. *Cognition & Instruction, 25*(2/3), 113-160.

Lunsford, M., Rowell, G., & Goodson-Espy, T. (2006). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. *Journal of Statistics Education, 14*(3).

Meletiou-Mavrotheris, M. (2003). Technological tools in the introductory statistics classroom: Effects on student understanding of inferential statistics. *International Journal of Computers for Mathematical Learning*, 8, 265-297.

Peters, S. (2011). Robust understanding of statistical variation. *Statistics Education Research Journal*, 10(1), 52-88. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ10(1)_Peters.pdf

Reaburn, R. (2011). *Students' understanding of statistical inference: Implications for teaching* (Doctoral dissertation, University of Tasmania). Retrieved from http://eprints.utas.edu.au/12499/1/Whole.pdf

Sanchez, E., & Inzunsa, S. (2006). *Meanings' construction about sampling distributions in a dynamic statistics environment*. Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematical thinking and learning* (pp. 465-494). New York: McMillan.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age Publishing.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*(2), 114-145.

Simon, M. (2013). *Developing theory for design of mathematical task sequences: Conceptual learning as abstraction.* Paper presented at the ICMI Study 22: Task Design in Mathematics Education.

Smith, T. (2008). *An investigation into student understanding of statistical hypothesis testing*. (Doctoral dissertation, University of Maryland, College Park). Retreived from Proquest. (3324893)

Steinbring, H. (1991). The theoretical nature of probability in the classroom. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 135-167). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publishing.

Watson, J. (2000). Preservice mathematics teachers' understanding of sampling: intuition or mathematics. Mathematics Teacher Education and Development, 2, 121-135. Retrieved from http://www.merga.net.au/documents/MTED_2_watson.pdf

PROSPECTIVE SECONDARY TEACHERS' CONCEPTIONS OF PROOF AND INTERPRETATIONS OF ARGUMENTS

AnnaMarie Conner, Richard T. Francisco, Carlos Nicolas Gomez, Ashley L. Suominen, & Hyejin Park
University of Georgia

We analyzed the interviews of three prospective secondary mathematics teachers to examine their conceptions of proof and how they validated arguments in the context of students' answers. Our participants had differing views of the definition of proof and its role in mathematics. Their work when validating arguments in large part aligned with their professed views of proof, with some deviations on the part of one participant. Further research must examine whether this consistency is prevalent across prospective teachers and how this relates to teachers' work with proof in classrooms.

Key words: Proof validation, Conceptions of proof, Prospective secondary teachers, Conviction

The role of proof in mathematics has been clearly established as significant. "Proving is one of the central characteristics of mathematical behavior and probably the one that most clearly distinguishes mathematical behavior from behavior in other disciplines" (Dreyfus, 1990, p. 126). Current national recommendations establish the desirability of elementary and secondary students engaging in reasoning and proof (National Council of Teachers of Mathematics, 2009; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Teachers' conceptions of proof, beliefs about the role of proof in mathematics, and their abilities to facilitate argumentation are related to how well they can implement these kinds of experiences (see, e.g., Conner, 2007). Little research has been devoted to how prospective secondary teachers develop and modify these conceptions in their university curricula. In this paper, we report results of a study in which we interviewed several prospective secondary mathematics teachers during their mathematics education coursework to examine their conceptions of proof and how they engaged in argument validation when arguments were situated in the context of student responses.

Relevant Literature

Teachers' conceptions of proof are inherently influenced by their experiences with proof in their mathematics coursework. Even though proof plays a central role in the undergraduate mathematics curriculum, numerous studies depict students' difficulties with proof production (e.g. Healy & Hoyles, 2000; Harel & Sowder, 1998). Students' lack of confidence with proof may be influenced by the fact that the field of mathematics cannot agree on a definition of proof (Hersh, 1993). However, even if students cannot give a formal definition of proof, many students have concept images of proof (Moore, 1994). Many studies have been conducted in which students at various levels were asked to construct proofs (see Reid, 2011), but as mathematics educators looked for more fine-grained explanations, some researchers have begun to examine students' validations of proofs (e.g., Knuth, 2002a; Selden & Selden, 2003; Weber, 2010).

Studies of proof validation have been conducted with various populations, including undergraduate students, practicing teachers, and research mathematicians. The results demonstrate that determining whether an argument is a valid proof is not straightforward. Selden and Selden (2003) asked undergraduate mathematics students whether given arguments proved a number theoretic statement. The aggregate of students' responses indicated a random response pattern. In another study, only six of thirteen undergraduate mathematics majors were able to determine that a real analysis proof was invalid (Weber & Alcock, 2005). A recent study on proof validation found that undergraduate students who completed an introduction to proofs course were often able to reject empirical arguments as proofs but again performed variably when asked whether a deductive argument (valid or invalid) was a proof (Weber, 2010). Research on practicing secondary teachers found that teachers accepted non-proof arguments as valid mathematical proofs (Knuth 2002a). Finally, Weber (2008) found that even practicing mathematicians do not always agree about whether an argument is a valid mathematical proof, even for relatively uncomplicated proofs (a couple of lines long). This ambiguity has important implications for teaching, as the final verdict of a proof's correctness is often determined by social norms (e.g. Hanna, 1991). It is therefore important to examine teachers' views of proofs, what they consider to be convincing, and how they validate arguments from students.

Theoretical Perspective

Our larger study coordinates a situative perspective on learning to teach mathematics (following Peressini, Borko, Romagnano, Knuth, & Willis, 2004) with current research on teachers' beliefs about teaching, mathematics, and proof (e.g., Cooney, Shealy, & Arvold, 1998; Ernest, 1993, 1988; Knuth, 2002a; Liljedahl, Rolka, & Rosken, 2007; Thompson, 1992). As we narrowed our focus for this particular part of the study, we coordinated several perspectives related to proof to provide guidance for our analysis.

The primary lens for our analysis of participants' conceptions of proof was the multiple roles that have been proposed for proof in mathematics. Proofs provide conviction that an assertion is true (e.g. Harel & Sowder, 1998) and justify mathematical assertions. De Villiers (1990) asserted that proofs play an important communicative role in mathematics and systematize the field. Other researchers have argued that proofs should also explain why an assertion is true (e.g. Hanna, 1990; Hersch, 1993). Following from these roles of proof in the discipline of mathematics, Knuth (2002b) contended that we must consider the following roles of proof in school mathematics: verification, explanation, communication, discovery, and systematization. In Knuth's (2002b) study, practicing secondary mathematics teachers reported some of these beliefs about the role of proof, including explaining why a statement is true, communicating mathematical knowledge, verifying the truth of a statement, and systematizing the field of mathematics, but lacked emphasis on promoting understanding. We examined what our participants viewed as roles of proof in mathematics and in the classroom.

An important goal for students in teacher education programs is the development of the ability to critically reflect upon students' thinking (Ball, 1988). One prominent way that mathematical knowledge is communicated is through written assignments and examination. Therefore, it is essential that teachers develop proficiency at reading mathematical proofs. We asked teachers to validate mathematical arguments (after Knuth, 2002a; Selden & Selden, 2003; Weber, 2010) by stating whether they qualify as mathematical proofs and whether they find them convincing. Our analysis of our participants' argument validations was informed by Selden and Selden's (2003) description of proof validation as a process by which someone reads and reflects on an argument in order to determine the extent to which it is correct. "Validation can include asking and answering questions, assenting to claims, constructing subproofs, remembering or finding and interpreting other theorems and definitions, complying with instructions (e.g., to consider or name something), and conscious (but probably nonverbal) feelings of rightness or wrongness" (Selden & Selden, 2003, p. 5). Because we were interested in participants' views of proof in the context of teaching mathematics, we situated our interview questions and proposed arguments as answers from hypothetical secondary students. The situative perspective was useful in making sense of their responses, as they often cited norms from undergraduate mathematics classrooms or referenced classroom teaching situations when giving their evaluations.

Methodology

This paper reports a subset of results of a larger study in which we followed sixteen prospective teachers through their mathematics education coursework. For this smaller study, we purposefully selected three prospective teachers and examined their perspectives on proof during their first year of mathematics education coursework. During this time, the prospective teachers were concurrently enrolled in mathematic courses that required regular engagement with proof (e.g., Abstract Algebra and Foundation of Geometry). Data collection for this study included three video-recorded semi-structured interviews of varying length (45 – 90 minutes). In the first interview, conducted during the first two weeks of the fall semester, participants were asked for their initial thoughts on the definition of proof, its role in mathematics, and its role in the mathematics classroom. In the second and third interviews, conducted at the end of the fall and spring semesters respectively, we asked students additional questions about proof and asked them to complete sets of proof validation tasks we had developed and adapted from other studies (see Table 1 for a summary of tasks). For example, several tasks asked students to read a set of arguments that purport to prove a particular statement and then decide whether or not it proved the statement. Some of the tasks were set in the context of a classroom in which different students had proposed the different arguments. Our protocol was based in part upon Knuth's (2002b) examination of practicing teachers' beliefs about the role of proof in mathematics and in their practice, with the argument validation tasks informed by other proof validation studies as well (e.g., Weber, 2010). Each interview was transcribed by a member of the research team and checked by another member to verify accuracy.

Table 1: Summary of Arguments Presented to Prospective Teachers

Problem/Claim	Argument	Argument Summary
Exponent Problem: Is it possible to select real values for *a* and *b* such that $(2^a + 1)^b$ would be an even number? Why or why not? (Interview 2)	Cathy's	2 to any power is even so $2^a + 1$ will always be odd. An odd number to any power is odd because if "foiled" the addition of one is consistent.
	David's	2 to any power is even so $2^a + 1$ will always be odd. An odd number to any power is odd because the last digit follows a cyclic pattern of odd numbers.
The **law of cosines** states that given ΔABC with sides of length *a*, *b*, and *c* respectively, then $c^2=a^2+b^2-2ab \cos C$ (Interview 2)	A	Pre-constructed dynamic geometry sketch, steps through a series of constructions, including a circle in which are two similar triangles. A chain of equations, written from proportional relationships, concludes the argument.
	B	Dynamic geometry sketch, user can move any vertex of the triangle ΔABC and observe measurements and calculations.
	C	Two cases using Pythagorean theorem: C is obtuse or C is acute.
	D	Distance formula on a coordinate plane with one vertex of triangle at (0, 0).
The **sum of the first *n* odd natural numbers** is n^2. ℕ = {1, 2, 3…} (Interview 3)	Bart's	Proof by example: First 10 cases shown.
	Daphne's	Visual proof with use of multi-colored dots in square arrays.
	Charlie's	Algebraic manipulation of sum of the first *n* odd

		natural numbers: $S(n) = 1 + 3 + \cdots + 2n - 1$.
	Eva's	Algebraic manipulation of summation formula of the first *n* natural numbers: $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$.
	Archie's	Proof by induction.
Number Theory Problem: For any positive integers *a* and *b*, if a + b is an odd number, then one *a* or *b* is an odd number and the other is an even number (Interview 3)	A	Argument has unnecessary algebraic manipulation to demonstrate that an even number plus one is an odd number.
	B	Proof by contradiction. Assumes *a* and *b* are even, finding an even sum. Then assumes *a* and *b* are odd, finding an even sum.
	C	Proof of converse.
	D	Three cases: *a* is odd and *b* is even, *a* and *b* are both odd, and *a* and *b* are both even.

To analyze the data we first coded the data and identified parts of the data in which participants talked about proof and proving in general, separating these from parts in which the participants were working on the proof tasks. Next, we summarized the participants' views of proof from their statements about proof and proving in general, and we summarized the participants' work with proof, paying attention to the characteristics of proof that our participants seemed to value. Our codes and themes were both analytic and inductive, as we began with knowledge of the purposes and characteristics of proof from the literature, but remained open to (and found) other purposes and descriptions mentioned by our participants.

Results

Our analysis of our data was guided by the following questions: What are the prospective teachers' views of proof and its role in mathematics? How do the prospective teachers analyze arguments from students? What are consistencies or inconsistencies in their talk about proof and analysis of students' arguments? Our participants had differing views of the definition of proof and its role in mathematics. Their work when verifying arguments in large part aligned with their professed views of proof, with some deviations on the part of one participant. In this section, we introduce Jill, Jason, and Vanessa, describe their views of proof, and briefly describe some of their proof validations.

Jill focused on issues of accuracy and being correct in both her general talk about proving and her examination of arguments. In interview 1, she described proving as "showing that it's correct and that it works." However, she does not believe she knows "the formal definition of proving" (Interview 1), implying that there is a correct formal definition. Jill believes that we prove things in math because otherwise we would just have to take someone's word for mathematical results, so we prove things to establish mathematical certainty:

> Well if we don't prove it and somebody just says hey, this is, this works, and then they don't prove it and then how do we ever know it really does work. Because if you just, you can take anyone's word for it, but if they don't prove it and show you why it works then you might never know if it's right or not. (Interview 1)

In her examination of arguments for various statements, Jill focused on examining the details of the various steps that were given. In particular, she examined the accuracy of the algebra within three of the arguments for the law of cosines, specifically questioning how the authors obtained various lines. She questioned a particular notation in Charlie's argument for the sum of the first *n* odd natural numbers, and she verified that she could see the differently sized squares in Daphne's argument for the sum of the first *n* odd natural numbers. She also referenced specific proof techniques or notations when she was talking about her own proving as well as examining students' arguments. For instance, she stated, "Because, when we are doing, like, proofs, and we have to talk about even and odd numbers, we would usually write $2x$ for an even number and then $2x$ plus 1 for an odd, for an odd number." She was uncomfortable with Charlie's argument for the sum of the first n odd natural numbers, saying, "they just went about it in an odd way." This argument seemed to be different from what she expected, and even though she concluded that it was a proof, she seemed to be looking for a trick of some sort that would make it not a proof.

When Jill talked about proof in the context of teaching and learning, she emphasized another aspect of proof: proof as a way to understand how and why something works. Jill's explanation is similar to that of Knuth's (2002b) participants who expected their students to learn "where statements come from or why they are true rather than accepting their truth as given" (p. 80). Her prime example of something to be proved is the quadratic formula:

> The quadratic formula to some kids is just like a bunch of letters, and they're like, "What do I do with these letters?" I don't get it. They just plug it in and it doesn't make, they are just like, "Okay, this is what I am doing. Plug it in, blah." They don't really understand what it, what's going on, but maybe if they proved it, they would see where those letters are coming from, where the numbers go in. (Interview 3)

However, when evaluating arguments, even arguments from students, she focused on the accuracy of the arguments, including their generality, their logical structure, and line-by-line analysis rather than the explanatory power of an argument.

Jason's conceptions of proof as illustrated by his answers to general questions about proof and proving and his examination of students' arguments were very consistent. Jason believes that proof and proving are integral parts of mathematics. He defined proving as "demonstrating why something is the case, not just saying that's the case. So you're building up your argument" (Interview 1). Jason explained that the roles of proofs in mathematics are verification, explanation, and logic outside of mathematics. In particular, he stressed the importance of proofs in relation to logic. When he analyzed the students' arguments, he pointed out what was being proved in their arguments, investigated if the arguments included all cases and examined each step of the arguments to determine if they made sense. He distinguished between illustrating a theorem and proving it when he analyzed the dynamic geometry argument for the law of cosines (argument B), Bart's argument for the sum of the first n odd natural numbers, and Cathy's solution to the exponent problem. In several cases, Jason criticized an argument for proving something other than the requested claim. This was true for three of the number theory arguments (A, B, and C).

When asked what students should prove, Jason focused more on the general concept of proving than on specific things to prove:

> I think they should come across the idea of proving something is true in all cases, that just proving that something works isn't the same as proving that something is always true. I think that's an excellent concept to teach students. (Interview 2)

In his proof validations, he tended to look for generality in an argument; for instance, he critiqued Bart's argument as not proving the claim in general. He also critiqued argument B for the law of cosines: "Technically you'd have to drag the cursor over an infinite amount of screen to prove it, so no, that's not proving it" (Interview 2, lines 540-542). Jason's view of the verification role of proof was illustrated by his answers to questions about how convincing the arguments were to him. In every case, Jason was either convinced by an argument and said it was a proof or was not convinced by an argument and said it was not a proof. This consistency was not observed in the other focus participants, and is contrary to the general trend of the findings of Segal (2000) and Weber (2010).

Vanessa's views of proof seemed to depend on her understanding of what a proof is or involves and how that coincided with the views of the instructor or the requirements of the course. Of the focus participants, Vanessa was the most accepting of arguments, including empirical arguments, as proofs. For instance, she accepted argument B for the law of cosines as a proof. In her examination of students' arguments, Vanessa did have some specific views about what a proof should look like. For instance, when examining David's solution to the exponent problem, she said that it was not what a formal proof should look like, but it made sense and was pretty convincing. She said that Archie's argument for the sum of the first n odd natural numbers was what she was used to seeing, so "I'm guessing" it's a proof (Interview 2). She critiqued Daphne's argument for the sum of the first n odd natural numbers as not a complete proof because she was used to "seeing a lot more writing and a lot more variables involved" (Interview 3).

Vanessa's definition of proof was flexible and considered the audience of the proof as an important factor, even at the beginning of her mathematics education coursework:

> To prove something is when…you're able to explain the concept or an idea to someone so that they can, like, understand it. It doesn't have to be ambiguous and like just mathematically jargon-filled and, like, complicated. It can be as simple as, like, a middle school person could understand it. So it's just a way for you to be able to explain something very well, so that somebody that it's not familiar with it can be able to really understand, I think. That's when you know that you've achieved the goal of proving something. (Interview 1)

When examining Eva's argument for the sum of the first n odd natural numbers, she essentially said that it was a proof for her but not for a high school student:

> But if, if I was like a high school student reading this. It wouldn't…make sense to me. It doesn't justify anything. Because I didn't know this fact [points to $n(n+1)/2$], so you're telling me to assume that fact, and then once I assume it then I should believe the rest. So to a high school student this is not a proof, this doesn't explain this statement right here, this claim right here. But to me, it makes sense as proof, because I know that [points to $n(n+1)/2$], and the whole thing just follows. (Interview 3)

When she examined Bart's argument in interview 3, she distinguished that it is a justification, which is appropriate for middle school, but it is not a proof. Vanessa's flexible definition of proof could be compared to Stylianides' (2007) definition of proof in K-12 mathematics, capturing the idea of considering classroom communities, even though she does not seem to acknowledge the deductive structure implied by Stylianides.

Implications for Future Research

If we want teachers to use proofs in ways that will promote students' understanding, we should provide opportunities for prospective teachers to consider the attributes of proofs and how they can be used to promote understanding. Our study shows that prospective secondary teachers validate students' arguments in ways that are consistent with the conceptions of proof they have developed during their school and university experiences. However, their developed conceptions seem to be individual, ranging from a flexible conception that is context-dependent and considers the audience to be an important factor, to a view that is focused on the accuracy and form of an argument, to a view that focuses on generality and logical structure of a proof. Each of these views of proof has aspects that would be useful to teachers of secondary mathematics, but each also contains aspects that could hinder teachers' assessments of student arguments. Future research should examine if teachers' validations of students' arguments remains consistent with their views of proof when larger numbers of participants are considered. In addition, research must examine what views of proof allow teachers to assist students in constructing and critiquing arguments in effective ways.

References

Ball, D. L. (1988). Unlearning to teach mathematics. *For the learning of mathematics, 8*(1), 40–48.

Conner, A. (2007). *Student teachers' conceptions of proof and facilitation of argumentation in secondary mathematics classrooms.* Unpublished Doctoral dissertation, The Pennsylvania State University, University Park, PA.

Cooney, T. J., Shealy, B. E., & Arvold, B (1998). Conceptualizing belief structures of preservice secondary mathematics teachers. *Journal for Research in Mathematics Education, 29*, 306–333.

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, *24*, 17–24.

Dreyfus, T. (1990). Advanced mathematical thinking. In P. Nesher, & J. Kilpatrick (Eds.), *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education* (pp.113-134). UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139013499.008

Ernest, P. (1988). The impact of beliefs on the teaching of mathematics. In P. Ernest (Ed.), *Mathematics Teaching: The State of the art* (pp. 249–254). London: Falmer. Retrieved June 12, 2008 from http://www.people.ex.ac.uk/PErnest/impact.htm.

Ernest, P. (1993). Constructivism, the psychology of learning and the nature of mathematics: Some critical issues. *Science and Education, 2*, 87–93.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6–13.

Hanna, G. (1991). Mathematical proof. In David Tall (Ed.), *Advanced mathematical thinking* (pp. 54–61). Dordrecht: Kluwer Academic Publishers.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Issues in Mathematics Education: Vol. 7. Research in collegiate mathematics education III* (pp. 234−283). Providence, RI: American Mathematical Society.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24*(4), 389–399.

Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33*(5), 379–405.

Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education, 5*(1), 61–88.

Liljedahl, P., Rolka, K., & Rosken, B. (2007). Affecting affect: The reeducation of preservice teachers' beliefs about mathematics and mathematics learning and teaching. In W. G. Martin, M. E. Strutchens, & P. C. Elliott (Eds.), *The Learning of Mathematics* (pp. 319–330). Reston: National Council of Teachers of Mathematics.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*(3), 249–266.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2009). *Focus in high school mathematics: Reasoning and sense making.* Reston, VA: Author.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics.* Washington, DC: Author. Retrieved June 1, 2012 from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Peressini, D., Borko, H., Romagnano, L., Knuth, E., & Willis, C. (2004), A conceptual framework for learning to teach secondary mathematics: A situative perspective. *Educational Studies in Mathematics, 56*(1), 67–96.

Reid, D. A. (with Knipping, C.). (2010). *Proof in mathematics education: Research, learning, and teaching.* Rotterdam, the Netherlands: Sense Publishers.

Segal, J. (2000). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior, 18,* 191–210.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38,* 289–321.

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teacher and learning* (pp. 127–146). New York: Macmillan.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*, 431–459.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning, 12*(4), 306–336.

Weber, K., & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics*, *25*(1), 34–38.

THE SELECTION AND USE OF EXAMPLES BY ALGEBRAISTS: AN EXPLORATORY STUDY

Abstract: This paper reports on an exploratory study of 10 algebraists designed to investigate the reasoning behind their selection of examples for their own teaching and research. Variation theory provided a lens with which to analyze the algebraists' goals for their collections of examples and to speculate about the resulting pedagogical implications. Though findings from this exploratory study should be regarded only as preliminary and in need of further justification, our results provide some initial evidence that mathematicians use a relatively small number of very well-chosen classes of examples in both their teaching and their research (suggesting that this might be a useful pedagogical strategy for students as well). We also report on the examples of groups and rings that the algebraists deemed to be the most important for students of introductory abstract algebra.

Key words: example usage, abstract algebra, variation theory

Introduction

Examples are believed to be very important in developing conceptual understanding of mathematical ideas (Mason & Watson, 2008; Tall & Bills, 1998; Tall & Vinner, 1981). Examples give insight into and can be used to create mathematical definitions, theorems and proofs (Cuoco, Goldenberg and Mark, 1997; Lakatos, 1976). Mathematicians use collections of examples to develop intuition and to generate, test, and refine conjectures (Michener, 1978). When a mathematician comes upon or creates a conjecture that is not obviously true, Courant (1981) claimed that the mathematician's first reaction is to call upon an example in order to think about the general through a particular case. The purposes of an example are to provide a more familiar and concrete means to explore ideas, and to evaluate constraints in theorem formulation. In particular, when used in such ways, they clearly indicate that the mathematician is attempting to work outside the symbolic system of the current problem.

There is great perceived pedagogical power of examples (Bills & Watson, 2008; Mason & Watson, 2008). Several studies have focused on student exemplification and use of examples in learning about concepts and proofs (Alcock & Inglis, 2008; Dahlberg & Housman, 1997; Mason & Watson, 2008). Moreover, studies have explored how graduate students use examples in determining the truth of conjectures (Alcock & Inglis, 2008). At this time, there are few studies of instructors' teaching with examples in undergraduate proof-based mathematics courses. For the purposes of this study, proof-based courses are undergraduate mathematics courses which focus on definitions, theorems and proofs.

This paper draws upon Watson and Mason's definition of an example as "any mathematical object from which it is expected to generalize" (2005, p. 3). In particular, we focus further on what Watson and Mason (2005) described as a "reference example"—that is, an example that is commonly used for testing conjectures and encodes all of the important information about the class of objects it exemplifies. This concept of a reference example suggests that, among the large space of examples of any particular concept that a mathematicians knows, there are some which are somehow more useful and called upon than others. We believe that most mathematicians have a very small class of reference examples that they repeatedly draw on in their work and explore whether this should have implications for their teaching practice. Moreover, while there is evidence that instructors do generally draw on a relatively small range of examples in their instruction, and that they present material to promote a reference example (e.g. the class $(\mathbb{Z}_n, +_n)$), this evidence is from a single case study (Fukawa-Connelly & Newton, in press), and there is a need to determine the generalizability of the results. Thus, the specific aims of this research are:

1) To explore the examples of groups and rings that algebraists think are most important for students to know and their reasoning for those choices.
2) To explore the truth of the hypothesis that mathematicians really only use a very small collection of examples, but incredibly well-chosen ones.
3) To draw tentative conclusions about the relationships between the responses to the above.

Pedagogical Uses of Examples

Researchers assert that "exemplification is a critical feature in all kinds of teaching, with all kinds of mathematical knowledge as an aim" (Bills & Watson, 2008, p. 77). To this end, research on example usage is on the rise. Several studies have focused on student exemplification and use of examples (Alcock & Inglis, 2008; Dahlberg & Housman, 1997; Mason & Watson, 2008). There are few studies documenting the use of examples by instructors in undergraduate proof-based mathematics courses. In proof-based courses, examples are often used in teaching to introduce a concept or motivate a definition; alternatively, they serve as a means by which students can attach meaning to definitions (Goldenberg & Mason, 2008). Goldenberg and Mason posited that, by exploring examples, "learners encounter nuances of meaning, variation in parameters and other aspects that can change" (2008, p. 184). That is, examples can be used to explore a definition (Fukawa-Connelly & Newton, in press), and, in such cases as Larsen and Zandieh (2008), create, test, and revise definitions.

Examples of concepts are also used by instructors in presentations of theorems and proofs. Mills (2012) identified a number of different ways instructors use examples in proof-based classes. Some of the types of examples that Mills identified can be thought of as giving insight into the creation of new mathematics. For example, she documented classes of examples that support motivating and exemplifying the statement of the claim. Similarly, Lakatos (1976) described the use of examples to articulate and refine definitions and conjectures. Such information can give insight into the creation of a proof or foster understanding of proofs (Alcock, 2010; Weber, 2010). Yet, in Mills' observations, instructors drew on these proof-supporting uses less than once per class period, and some only a few times across the entire semester. Moreover, she did not describe what examples the instructors use or how they think through their choices. Thus, there is still need to investigate how mathematicians think about example use both in their mathematics and in their teaching.

Theory

Variation theory is a natural means to analyze the affordances offered by the collection of examples that students experience and has been used to describe goals for presenting examples (Goldenberg & Mason, 2008). Variation theory suggests that the way a person develops an understanding of a concept depends on which aspects the individual can discern (Runesson, 2006). In order for a person to be able to discern a particular aspect of a concept, he or she must experience variation of that aspect (Runesson, 2006). In particular, if the instructor's proposed examples all have similar characteristics, there will be significant variation that students will not experience.

Tall and Bills' (1998) asserted that a definition-theorem-proof approach to proof-based courses will generally be unsuccessful without also giving students the ability to develop a rich intuition from their experiences. Tall also claimed the importance of examples that vary in complexity so that students are able to apprehend important features. Thus, this study focuses on examples of concepts because they are uniquely powerful in both mathematics and the teaching of mathematics, and seeks to explore Tall's suggestions about how mathematicians approach this idea of rich images of concepts in their work and teaching.

Methods

Because we are interested in exploring the relationship between algebra instructors' use of examples in their own mathematical work and teaching, we constructed a purposeful sample of algebraists who also teach undergraduate algebra classes. Our survey was specifically designed to assess (1) which examples algebraists considered most relevant to their research and (2) which among those examples they considered most important for students in an introductory abstract algebra class. We analyzed data by borrowing some techniques from grounded theory while keeping extant literature about instructors' demonstrated pedagogical example use in mind (Fukawa-Connelly & Newton, in press; Mills, 2012). We began by repeatedly reading responses to the survey, making summative comments about the categories of examples that algebraists claimed were important, and characterizing those important to their research as reference examples. We then analyzed the characteristics of the examples deemed pedagogically important via the method proposed in (Fukawa-Connelly & Newton, in press) and then aggregated across the group of mathematicians. Finally, we attempted to discern if individual mathematicians generally use a very small class of examples.

Data and Results

1) Examples of groups and rings that algebraists believe introductory students should know.

Groups: Algebraists believe that the symmetric groups of small order are the most important class of groups for students to be familiar with as they all mentioned at least one, followed by the cyclic groups, which were cited by seven (one of the algebraists listed the integers, six did not list any specific cyclic group, nor did they single out the cyclic groups of prime order). The dihedral groups were the next most mentioned with six (naming the general class and one listing a particular example), followed by the matrix groups (typically over a finite field). Three mentioned direct products as a means of creating other types of groups (the lattice groups and finite abelian groups, while one stated "direct products of groups"). One participant mentioned elliptic curves, continuous functions, and the orthogonal groups. The dihedral groups were specifically described as an ideal introductory example because of their focus on the behavior of the group structure.

Rings: There is less agreement among the algebraists about the full range of examples of rings students should experience, but there were still commonalities amongst their responses. The most commonly cited example was the matrix rings (by seven participants), followed by polynomial rings (six participants), then the integers (five participants), and the cyclic rings (four participants). The real numbers were listed by two participants and no other examples were listed more than once, but 10 different examples of classes were listed once, including the class of "fields" as an example that students should be familiar with at the end of a first semester course.

The algebraists primarily gave two explanations for their choices. First, they asserted that the given examples are easy to work with, making these examples ideal for performing calculations and verifying results. Second, the algebraists chose the given examples to include structures exhibiting varying mathematical structure (with regards to, for example, commutativity, finiteness, inclusion of zero-divisors), in order that students be exposed to a variety of structures that behave differently (and perhaps non-intuitively). The following explanation, while focused on individual examples, demonstrates most of the types of reasoning that the mathematicians drew upon:

> For groups, knowing basic properties of cyclic groups, S_3 and D_8 will get you through quite a bit. Knowing about S_4 and A_4 definitely expands your range substantially. If a student knows these along with the quaternion group, they probably can test just about anything reasonable at the first semester of study level.

That is, the algebraists' thinking about examples appear to implicitly reflect the implications of variation theory (Runesson, 2006).

There were a small number of examples whose absence we deemed notable, especially given the rationale that went into the algebraists' selection of examples. Finite fields, while being mentioned in conjunction with matrix groups (over a finite field), were not directly mentioned in their own right. They were indirectly mentioned, of course, with each reference to the modular rings $\mathbb{Z}_n$(and perhaps the algebraists believed this to be sufficient). This is noteworthy not only because of the significance of finite fields in field theory and number theory but also because they seem to fit the algebraists' criteria perfectly: finite fields are small, easy to perform calculations with, and are certainly a unique example of varying structure. The second exclusion was the ring of quaternions (though, curiously, the quaternion group was referenced), being the only accessible example of a division ring. It is possible that this example was not included because division rings do not receive much attention in an introductory abstract algebra course, yet we still found it surprising that a paradigmatic example of an accessible structure exhibiting deviant behavior received no attention.

2) *Algebraists tend to think about classes of objects as examples, especially pedagogically.*

While there were a small number of individual examples of groups and rings listed (such as the integers, rationals, quaternions or specific permutation and dihedral groups), all of the surveyed algebraists listed at least one class of groups and at least one class of rings (most listed multiple). Five of the surveyed algebraists listed only classes of groups as their examples. The pattern was similar with rings, although half listed the integers and more of the participants had a single example of a particular class that they listed as exemplary (such as $\mathbb{Z}[\sqrt{-5}]$). As a result, while all listed at least one class of examples, only one listed classes of ring examples exclusively.

This finding suggests that Fukawa-Connelly and Newton's (in press) observation that instructors quickly transition from discussing individual examples to classes of examples is likely to be common. That is, if the instructors think primarily of classes of examples, then any specific example would be understood as essentially interchangeable with any other members of the class, and, therefore, it would be quite reasonable to talk about a "general" element, such as $\mathbb{Z}/n\mathbb{Z}$ or S_n almost immediately. Then, as the professor in Fukawa-Connelly and Newton's work did, theorems about all members of the class might be proven.

What we know about learning suggests this is somewhat incompatible with how students think, at least as they are learning new concepts. In particular, students are not likely to think of a broad class of examples right away. Rather, students are likely to think about quite specific individual cases and then take time to apprehend how they all have similar characteristics (Wagner, 2010). In this way, this process of moving from thinking about individual characteristics to seeing structural similarities is difficult for students and requires significant time, even when the common structure in the presented examples are transparent to mathematicians (Dubinsky, Dautermann, Leron, & Zazkis, 1997).

3) *Most algebraists draw on relatively small collections of examples for most of their work.*

The mathematicians were asked what five examples they most commonly use in their research. One algebraist's research was about a particular type of Lie algebra, and, as a result, found that the question was not meaningful. Another participant acknowledged that he studied the sporadic simple groups, quipping that these included 26 examples (rather than 5). There were also a number of other more abstract (esoteric) structures listed, normally by one particular participant, suggesting specificity to that particular research area. These responses were not representative overall, however, as many listed an array of examples or classes of examples. One explained that the groups he most frequently used included the "dihedral groups, quaternion group, cyclic groups, symmetric groups, matrix groups." Many provided similar responses; in general, these algebraists generally used a relatively small number of classes of examples.

The participants also indicated that the listed classes of examples generally constitute a large fraction of the examples they use, suggesting that these examples operate as reference examples. One person did not respond to the question and one (who listed the sporadic simple groups) remarked that, because there are an infinite number of simple groups, the examples he listed represented 0% of those he used in his research. The other mathematicians claimed that the five listed examples accounted for 50%, 70% ,75%, 75%, 80%, 80%, 90% and 100% of the examples used in their work. This data generally reflects that most of their work is done with very few examples (or classes of examples). Although this recommendation should not be taken as dispositive, the data suggests that it might be more useful for students to have deep familiarity with a few examples, rather than a more broad experience with many. That is, mathematicians appear to use few examples, but think deeply about them, and there is reason to believe that this type of practice may be equally valuable for undergraduates (Collins, Brown, & Newman, 1989). Mathematicians have significant experience with and success at mathematics learning, and continue to learn new mathematics. As such, their habits suggest potential strategies for undergraduates (Collins, et al., 1989). This stands in direct contrast to recent studies that have either explicitly or implicitly suggested that undergraduates should have experience with a large number of examples (Watson & Mason, 2005; Sinclair, Watson, Zazkis, & Mason, 2011).

The algebraists primarily cited two reasons for their choices of examples in their own research. In agreement with Courant (1981), the most commonly cited was the use of examples for illustrative purposes – that is, to make the general more concrete (six participants). This was followed by those who mentioned using examples as building blocks to aid in generalizing definitions and theorems – that is, making the concrete more general (four participants). It is curious to note that only one mathematician cited the testing and proving of conjectures; the reasons for this omission are presently unclear.

Discussion

While it is inappropriate to draw major conclusions from a small-scale study, we do cautiously summarize our findings and interpret them in terms of their implications for the field. First, it is important to note that this is meant as an exploratory study and the hypotheses that we generated will be subsequently tested via a large-scale online survey (c.f., Lai, Weber, & Mejia-Ramos, 2012). We see such a survey as a logical extension of our work and as a means to generalize from prior case studies (e.g., Fukawa-Connelly & Newton, in press).

Our data suggests that algebraists generally believe that students need to be proficient with a relatively small number of classes of examples of groups and rings, with few mentions of individual members of those classes ($\mathbb{Z}$, S_3, A_4, and D_8 being exceptions). Our data collection was limited in that we did not ask instructors to describe useful structures that do not fulfill all of the properties of a given structure. None of them listed any, so we cannot know the reasons for this; it might be because we did not ask, because they do not consider them important, or for some other reason. Moreover, we wonder about the pedagogical implications of their listing of classes of examples. Fukawa-Connelly and Newton (in press) showed one example of a professor quickly transitioning from introducing a specific example from the class of $\mathbb{Z}_n$ to proving theorems about the entire class (one lecture after introducing the first example). We suggest that there are significant implications for student learning about the specific content and also what such practices communicate to students about what is valued in mathematics and mathematics classes (if you can't keep up, you're not a good math student).

Second, we see a distinction between how mathematicians use examples and how they expect students to use them. Their examples for students seem to be geared towards exposure to certain notions; there's a certain "expanding of their horizons" tone to some of the rationale. For their own work, the algebraists focus on using examples to illustrate and generalize claims. While there is some overlap in those contexts, perhaps one of the limitations is our failure to ask the algebraists to explain how they show students how to ensure appropriate boundaries for generalizations. For example, as noted above, we did not ask for near-examples which would give indications about structures that would allow students to test the boundaries of generalizations. Moreover, while none of the mathematicians specifically described using any of their listed structures as counterexamples, it seems reasonable to suggest that they see determining the limits of generalization via counterexamples as part of the generalization process (Lakatos, 1976; Zazkis & Leiken, 2008). From a research perspective, we will next revise our data collection tool to both give better access to algebraists' thinking and to perform appropriately in an online setting.

References

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Holton & P.W. Thompson (Eds.), *Research in collegiate mathematics education VII* (pp. 63-91). Washington, DC: MAA.

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69*, 111-129.

Bills, L., & Watson, A. (2008). Special issue: Role and use of exemplification in mathematics education [Editorial introduction]. *Educational Studies in Mathematics, 69*, 77-79.

Collins, A., Brown, J.S. & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. In L.B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Courant, R. (1981). Reminiscences from Hilbert's Gottingen. *Mathematical Intelligencer*, *3*(4), 154-164.

Cuoco, A., Goldenberg, E. P., and Mark, J. (1997). "Habits of Mind: an organizing principle for mathematics curriculum." *Journal of Mathematical Behavior*, *15*(4), 375-402.

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics, 33*, 283-299. doi:10.1023/A:1002999415887

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics, 27*, 267-305.

Fukawa-Connelly, T., & Newton, C. (*in press*) Evaluating mathematical quality of instruction in advanced mathematics courses by examining the enacted example space. *Educational Studies in Mathematics.*

Goldenberg, P., & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics, 69*, 183-194. doi: 10.1007/s10649-008-9143-3

Lai, Y., Weber, K., & Mejia-Ramos, J. P. (2012). Mathematicians' perspectives on features of a good pedagogical proof. *Cognition and Instruction, 30*(2), 146-169.

Lakatos, I. (1976). *Proofs and refutations.* Cambridge: Cambridge University Press.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67*, 205-216.

Mason, J., & Watson, A. (2008). Mathematics as a constructive activity: Exploiting dimensions of possible variation. In M. Carlson, & C. Rasmussen (Eds.), *Making the connection:*

Research and practice in undergraduate mathematics (pp. 189-202). Washington, DC: MAA.

Michener, E. R. (1978). Understanding understanding mathematics. *Cognitive Science, 2*, 361–383.

Mills, M. (2012, February). Investigating the Teaching Practices of Professors When Presenting Proofs: The Use of Examples. Conference on Research in Undergraduate Mathematics Education, Portland, OR.

Runesson, U. (2006). What is it possible to learn? On variation as a necessary condition for learning. *Scandinavian Journal of Educational Research*, *50*, 397-410.

Sinclair, N., Watson, A., Zazkis, R., & Mason, J. (2011). The structuring of personal example spaces. *The Journal of Mathematical Behavior*, *30*, 291–303

Tall, D. & Bills, L. (1998). Operable Definitions in Advanced Mathematics: The case of the Least Upper Bound,*Proceedings of PME 22*, Stellenbosch, South Africa, 2, 104– 111

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics, with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151-169.

Wagner, J. (2010). Transfer-in-pieces. *Cognition and Instruction, 24*(1), 1-71

Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Mahwah, NJ: Erlbaum.

Weber, K. (2010, February). *The pedagogical practice of mathematicians: Proof presentation in advanced mathematics courses*. Address at the Conference on Research in Undergraduate Mathematics Education, Raleigh, NC. Retrieved from http://sigmaa.maa.org/rume/crume2010/Archive/Weber_Keith.doc

Zazkis, R., & Leikin, R. (2008). Exemplifying definitions: A case of a square. *Educational Studies in Mathematics*, *69*(2), 131-148.

47

LIVING IT UP IN THE FORMAL WORLD: AN ABSTRACT ALGEBRAIST'S TEACHING JOURNEY

John Paul Cook — University of Science and Arts of Oklahoma
Ameya Pitale, Ralf Schmidt, Sepideh Stewart — University of Oklahoma

Abstract algebra is a fascinating field of study among mathematics topics. Despite its importance, very little research has focused on the teaching of abstract algebra. In response to this deficiency, in this study we present an abstract algebra professor's daily activities and thought processes as shared through his teaching diaries with a team of two mathematics educators and another abstract algebraist over the period of two semesters. We examined how he was able to live in the formal world of mathematical thinking while also dealing with the many pedagogical challenges that were set before him during the lectures.

Keywords: Reflections on Teaching, Abstract Algebra, Formal World of Mathematical Thinking

Introduction

William Thurston (1994), the Fields medalist, posed the question: "How do mathematicians advance human understanding of mathematics?" In his view, "what we are doing is finding ways for people to understand and think about mathematics" (p.162). It is unclear to what extent this is manifested in teaching practices at the undergraduate level, however, which are largely undocumented in the literature. As Dreyfus (1991) suggested, "one place to look for ideas on how to find ways to improve students' understandings is the mind of the working mathematician. Not much has been written on how mathematicians actually work" (p. 29).

This statement is still relevant almost two decades later, as Speer, Smith, and Horvath (2010) believed that very little research has focused directly on teaching practice and what teachers do and think daily, in class and out, as they perform their teaching work. They found that often "strong content knowledge and the ability to structure it for students may be taken as sufficient for good teaching" (p. 111).

Research in pedagogy at the university level is fairly new, and regrettably the communication between the mathematicians and those outside of the community is often very limited. According to Byers (2007):

> People want to talk about mathematics but they don't. They don't know how. Perhaps they don't have the language, perhaps there are other reasons. Many mathematicians usually don't talk about mathematics because talking is not their thing – their thing is "doing" of mathematics. (p. 7)

In an attempt to close this gap, Hodgson (2012), in his plenary lecture at ICME 12, raised the point about the need for a community and forum where mathematicians and mathematics educators can work as closely as possible on teaching and learning mathematics. In recent years various institutes and individuals have been more willing to examine and reflect on their own teaching styles, leading to a growing body of research in this area. For example, a study by Paterson, Thomas and Taylor (2011) described a supportive and positive association of two groups of mathematicians and mathematics educators from the same university which allowed the "cross-fertilization of ideas" (p. 359). The group met on a regular basis and discussed teaching strategies while watching small clips of each other's videos during a teaching episode. Hannah, Stewart and Thomas (2011, 2013) indicated cases in which a mathematician took careful diaries of his/her actions and thoughts during linear algebra lectures and reflected on them with the rest of the team. Also, a study by Kensington-Miller, Yoon, Sneddon and Stewart (2013) showed how a mathematician, with the help of mathematics educators in a research team, made changes in his lecturing style while teaching a large undergraduate mathematics course by asking well-planned questions.

Naturally, the teacher and any corresponding methods of instruction do not stand alone in the classroom; they are subject to the needs and abilities of the students. In addition to a lack of information about teaching practices in abstract algebra, there is considerable evidence documenting student difficulty with the subject's most basic concepts (Clark, Hemenway, St. John, Tolias, & Vakil, 2007; Dubinsky, Dautermann, Leron, & Zazkis, 1994). This situation has led one group of researchers to starkly conclude that "the teaching of abstract algebra is a disaster" (Leron & Dubinsky, 1995, p. 227). To further investigate what makes this course so challenging, we will examine an abstract algebraist's daily mathematical activities through his teaching diaries to understand his way of thinking and possible challenges of teaching advanced mathematics courses that many mathematicians (and their students) may face. The overarching aim of this study is to investigate how mathematicians live and dwell in the formal world of mathematical thinking and, at the same time, communicate their knowledge to their students. Our research questions are: Given that the mathematician in this study is a formal thinker, how does he invite students to his world and to what extent is he willing to help students to reach the higher level of mathematical thinking?

In the next section we will explore a theoretical framework by Tall (2004, 2010, 2013) that is appropriate in guiding this research to help us understand more about mathematicians as formal thinkers.

Theoretical Framework

In his theory, Tall introduced a framework based on three worlds of mathematical thinking: the *conceptual embodiment*, *operational symbolism* and *axiomatic formalism*. The world of conceptual embodiment is based on "our operation as biological creatures, with gestures that convey meaning, perception of objects that recognise properties and patterns...and other forms of figures and diagrams" (Tall, 2010, p. 22). Embodiment can also be perceived as the construction of complex ideas from sensory experiences, giving body to an abstract idea. The world of operational symbolism is the world of practicing sequences of actions which can be achieved effortlessly and accurately. The world of axiomatic formalism "builds from lists of axioms expressed formally through sequences of theorems proved deductively with the intention of building a coherent formal knowledge structure" (p. 22). Tall (2013) suggested that:

> Formal mathematics is more powerful than the mathematics of embodiment and symbolism, which are constrained by the context in which the mathematics is used. Formal mathematics is future-proofed in the sense that any system met in the future that satisfies the definitions of a given axiomatic structure will also satisfy all the theorems proved in that structure. (p. 138)

In his view "research mathematicians will focus attention on the higher demands of research and assert professional standards appropriate at that level" (p. 143). As these levels are decidedly relevant in abstract algebra, we employed this framework as a means of differentiating and drawing comparisons between the varying levels of mathematical thinking exhibited by the mathematician in his journals.

Method

The research described here is a case study of a research mathematician that took place at a large research university in Fall 2012 and Spring 2013. The research team consisted of two mathematicians and two mathematics educators forming a community of enquiry. The data for this research comes from one of the research mathematicians' daily reflections on his teaching of an abstract algebra course, which were made available to the group after each class; the team members' observation of the classes and their comments; weekly discussion meetings of the whole group after reading each of these reflections and the audio recordings of each meeting which were later transcribed.

The mathematics professor in this study was an experienced faculty member who had taught many mathematics courses from college algebra to algebraic geometry. He captured many details in his daily diaries and shared them promptly with the rest of the research team. The journals were brief, often included technical language, and gave an impression to the reader of being present in the class. During the weekly meetings, the rest of the research team, having already read the journals, gave the mathematician an opportunity to discuss his teaching from the past week. This was followed by questions from the research team, which often generated additional discussions. Additionally, the mathematician welcomed unannounced visits to his class by other members of the team. During the course of the two semesters he planned and devised a few teaching experiments in his abstract algebra lectures. He was approachable and open to new ideas during the meetings, and even attended educational talks by graduate students in the mathematics department. His positive attitude toward teaching and education enabled the team to get as close as possible to his way of thinking and interacting in the classroom.

The main themes emerging from the data were: the role of questions during the lectures; the role of examples to preview and illustrate a concept; assessment of students, content, class as a whole and students' understanding; examining the content, textbook and homework; reflection on himself, teaching, preparation, interaction and content; teaching experiments, preparation, decision making before/during the class, philosophy, rapport, teaching observation based on experience and teaching details. The data related to teaching and reflection comprised half of the data. For the purpose of this paper we will only concentrate on teaching experience and reflections on teaching.

What we will illustrate next is a glimpse into the mathematician's daily activity as a formal thinker stepping out of the research world and entering into the classroom to teach.

Results

As mentioned above, the mathematician performed several teaching experiments during the course of two semesters. The teaching experiments often included more focus on students and less lecturing. In one occasion he wrote in his diary (April 1):

> *The second half of the class was spent on the notion of conjugacy classes. I did not do this at all the way I had prepared it. Somehow, the idea of me writing "Definition:..." etc seemed really boring. Instead, I decided to introduce the notion by means of the simplest non-trivial example, namely S_3. By now everyone is familiar with this group. I resorted to the trick I had used before of having not me, but a student write on the blackboard. I called for a volunteer; nobody was eager, but eventually one of the better students stood up. ...It is always interesting how the simple fact that it is not me but a student standing there seems to increase class participation immediately. I believe this example was very illustrative and they learned the notion of conjugacy class better than with a formal definition. I am now wondering whether to even do the formal definition at all next time we continue, or just leave it at that.*

Here we see a formal thinker moving away from his comfort zone of definitions, theorems and proofs, reverting back to a simple example in the symbolic world of mathematical thinking. By involving a student and starting from an example he is attempting to reach out to his students and break down the abstraction of this concept. Suggesting that he viewed this break from the routine as successful, he mentioned that he is considering not making use of the formal definition at all the next time he teaches the course.

On another occasion (February 11) before the class he decided to invite students to construct their own proofs without writing it all on the board for them:

In this class we proved the main theorem of Galois theory. Looking at my notes right before class, I realized that all the pieces of the proof are in place, and there isn't really much more to do. So I decided to more or less have the students develop the proof. I started the class by not stating the theorem, but writing down some ingredients of the proof, without the students knowing that this is going to be a proof at all. Then I guided them towards the main theorem by asking questions. Within ten minutes the proof was complete, and only then did we state it formally as a theorem.

It was apparent that as an experienced mathematician he knew his material well but was consciously aware that his students were not yet at that level. This awareness of the discrepancy between his understanding and that of his students was a common theme throughout his journals. For example, on September 17, he wrote:

We spent the last 15 minutes proving that a field of fractions is a fraction field. I am afraid that the point of this was not entirely clear, and that it was in fact a little bit confusing. But I don't know how to do it better.

It was also clear that he was looking for ways to make the ideas that came so naturally to him more accessible to his students (February 13):

Then we formulated and proved a small Galois-theoretic result. This would have been kind of a boring afterthought to the main theorem, with no obvious immediate purpose. So I thought to myself right before class, how can I make this interesting? I resorted to the following trick: Before stating the theorem, I said that the level of complexity of the proof is such that a similar statement could easily be a problem on a qualifying exam. I think this kept everyone on their toes for the duration of the proof. I halfway had the class develop the proof, and it seemed like everyone was thinking hard, wanting to prove to themselves that they would be able to figure out a qualifying exam problem.

He knew that performing these teaching experiments would come with a cost, so he was consciously aware of the time and often was battling a tension between his identity as a mathematician (and the desire for conciseness and formality) with his identity as an instructor (and the desire to break down the material for his students) (October 22):

How did this happen? For one, I really wanted to get through with the proof of Gauss' theorem today. I knew time would be tight, and indeed we barely made it. So from the beginning I was in lecturing mode. Almost as if I didn't want the class to be disturbed by the possibility of students asking questions. This, of course, is a terrible attitude towards teaching.

As he reflected on his teaching, this conflict continued to be a challenge. After performing another teaching experiment he wrote: "While this was a very 'cool' and constructive class, we made zero progress on the material we are supposed to cover in this course" (November 2). Though teaching experiments comprised a very small portion of the course overall, they provided a rich source of insight into this mathematician's efforts to help his students navigate the formal nature of abstract algebra.

Although the professor was happy to try different teaching methods, at times he was not ready to change his beliefs on the usefulness of traditional lectures (October 15):

…I thought back about my own algebra education today, and how this was all very old-fashioned classroom lectures. There were not many questions asked by the teacher, and certainly there was never any group work or any kind of teaching experiment. Nevertheless, I remember thoroughly enjoying every class, with most of the fun coming from the beauty of the material itself. Made me wonder if it was just me having fun, or if there is more value in old-fashioned lecturing than we usually think there is.

The weekly discussions with the rest of the research team gave the mathematician another opportunity to reflect on his teaching and speak freely about the past week's events. It was noted that he was often excited about teaching, especially his favorite concepts and the materials that were well-prepared before the class (September 15):

> *…So when I prepare a class then it's also understanding for myself even though its stuff I know in principle but I put it fresh on my mind right, and it's almost like it wants to come out and it's really often that I wish the class would happen right now because I want to tell people about it now…*

Concluding Remarks

For a research mathematician, transitioning from the formal world of mathematical thinking back to the symbolic and embodied worlds is pedagogically challenging and requires an awareness of students' level of thinking and careful preparation. The results in this paper give a brief account of the mathematician's everyday actions and thoughts. Returning to our research questions, the results of this study provide some insight into the thought processes engaged in by a mathematician teaching an abstract algebra course. Specifically, this paper details his reflections on his efforts to help his students access the formal nature of abstract algebra. The results of this study provide a preliminary characterization of his efforts to do so. Of course, this is but a small portion of his journals and reflections (a full-scale report is beyond the scope of this proposal). The authors are in the process of making the full report of this research available in the near future.

Reflecting on the statement that "the teaching of abstract algebra is a disaster," the results of this two-semester study of an abstract algebra course suggest a positive outcome with regards to the collaboration between mathematicians and mathematics educators, despite the fact that 95% of the course consisted of old-fashioned blackboard lectures and no classroom technology was deployed. So far the effect of this collaboration has been positive in the sense that everyone in the group are not only focusing on the research mathematician's teaching strategies and thinking processes, but also their own teaching and decision making on a day-to-day basis. Moreover, it has provided a platform allowing mathematicians to talk about mathematics freely and share their pedagogical challenges with each other.

References

Byers, W. (2007) *How mathematicians think: Using ambiguity, contradiction, and paradox to create mathematics*, Princeton University press.

Clark, J., Hemenway, C., St. John, D., Tolias, G., and Vakil, R. (2007). Student attitudes toward abstract algebra, *Primus*, *9*(1), 76-96.

Dreyfus, T. (1991). Advanced Mathematical thinking processes. In D. O. Tall (ed.) *Advanced Mathematical Thinking*, (pp. 25-41). Dordrecht: Kluwer.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory, *Educational Studies in Mathematics, 27*(3), 267-305.

Hannah, J., Stewart, S., & Thomas, M. O. J. (2011). Analysing lecturer practice: the role of orientations and goals. *International Journal of Mathematical Education in Science and Technology*, *42*(7), 975-984.

Hannah, J., Stewart, S., & Thomas, M. O. J. (2013). Conflicting goals and decision making: the deliberations of a new lecturer, In Lindmeier, A. M. & Heinze, A. (Eds.). *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 2, pp. 425-432. Kiel, Germany: PME.

Hodgson, B. R. (2012). Whither the mathematics/didactics interconnection? Evolution and challenges of a kaleidoscopic relationship as seen from an ICME perspective. ICME 12 conference, Plenary Presentation, Seoul, South Korea.

Kensington-Miller, B., Yoon, C. , Sneddon, J. Stewart, S. (In press). Changing beliefs about teaching in large undergraduate mathematics classes. *Mathematics Teacher Education and Development.*

Leron, U. & Dubinsky, E. (1995). An abstract algebra story, American Mathematical Monthly, *102*(3), 227-242.
Paterson, J., Thomas, M. O. J., & Taylor, S. (2011). Decisions, decisions, decisions: What determines the path taken in lectures*? International Journal of Mathematical Education in Science and Technology, 42*(7), 985-995.
Speer, N. M., Smith, J. P & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, *29*, 99–114.
Tall, D. O. (2004). Building theories: The three worlds of mathematics. *For the Learning of Mathematics*, *24*(1): 29-32.
Tall, D.O. (2010). Perceptions, operations, and proof in undergraduate mathematics. *Community for Undergraduate Learning in the Mathematical Sciences (CULMS) Newsletter*, 2, 21-28.
Tall, D. O. (2013). *How humans earn to think mathematically: Exploring the three worlds of mathematics*, Cambridge University Press.
Thurston, W. P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, *30*(2), 161–177

PRESENTATION OF MATRIX MULTIPLICATION
IN INTRODUCTORY LINEAR ALGEBRA TEXTBOOKS

Abstract: We conducted an analysis of 17 modern, introductory linear algebra textbooks to investigate presentations of matrix multiplication. Using Harel's (1987) textbook analysis framework, we examined the sequencing of matrix multiplication and its accompanying rationale. We found two principal sequences: one which first defines the operation as a linear combination of column vectors before introducing the dot product method (LC to DP), and another which invokes the dot product method before linear combinations (DP to LC). The rationale for these two trajectories varied in interesting ways. LC to DP demonstrates that solving a system of linear equations is equivalent to solving its corresponding matrix equation Ax=b. The rationale for DP to LC was less focused, opting in several cases to postpone the explanation until linear transformations are covered. We hope to initiate a discussion about the effectiveness of and pedagogical implications for these two contrasting approaches.

Key words: linear algebra, matrix multiplication, textbook analysis

Introduction

Matrix multiplication is likely the first abstract multiplication that students encounter in undergraduate mathematics. It is likely to be the first multiplication that does not 'multiply' in the literal sense (as with scalar multiplication or the multiplication of real numbers). Rather, matrix multiplication is a multiplication (in the sense of ring theory) because it is associative and distributes over matrix addition. As such, it seems reasonable to expect some hesitancy from students to accept this more abstract operation (even though the computations are relatively straightforward). This elicits important questions. Since matrix multiplication can be defined in many different ways (see, for example, Carlson, 1993), how is it being taught in undergraduate classrooms? How is it being explained and motivated?

While no studies were found directly examining teaching practices of this particular topic in linear algebra, a possible avenue of potential insight is to investigate presentation of matrix multiplication in linear algebra textbooks. Harel (1987) presented an analysis of linear algebra textbooks, yet our work is distinct in two important ways. First, Harel's analysis was nearly three decades ago, a significant period of time in which impactful attempts at linear algebra curriculum reform have been made (for example, Carlson, Johnson, Lay, & Porter, 1993) and an array of new textbooks have been published. Second, Harel makes no direct mention of how matrix multiplication is defined or explained. Harel's findings, however, provide a useful framework with which to conduct our analysis. He found that linear algebra textbooks varied on the basis of sequencing of content, generality of vector space models, introductory material, embodiment, and symbolization. Those tenets of Harel's framework that inform our analysis are detailed in the next section.

This paper seeks to use Harel's (1987) framework to investigate the presentation of matrix multiplication in modern, introductory linear algebra textbooks. In doing so, we sought answers to the following research questions:

- How is matrix multiplication defined in modern textbooks?
- What rationale is given for the proposed definition(s)?
- What are the pedagogical implications of any differing approaches?

Theoretical Framework

We employ Harel's (1987) framework for textbook analysis. However, as Harel's paper presented a macro-analysis (of the content presentation on a general scale throughout entire textbooks) and this paper presents a microanalysis (of the presentation of one specific topic), we adapted the framework to fit the parameters of this study. Those tenets relevant to our very specific analysis of matrix multiplication are sequencing of content and introductory material. We restrict ourselves to these two to form the basis of our analysis.

Sequencing of content. Harel noted that introductory textbooks typically follow a computation-to-abstraction approach, in which systems of equations and matrix multiplication are used to necessitate vector spaces and more general mathematical structure. Restricting our focus specifically to matrix multiplication, however, we examined the sequencing in which the textbook authors proceed with the different methods of performing this operation. Our anecdotal evidence and experience indicate that each textbook contains at least two methods. What is less clear, however, is the order in which these methods appear and the rationale given for that particular approach. Any trends in this regard would provide insight not only into the overall pedagogical philosophies employed in these textbooks, but would also provide preliminary indications of how this topic is being taught in undergraduate classrooms.

Introductory material. Harel found that introductory material, attempting to bridge the intellectual gap between prior knowledge and the new mathematics to be learned, was presented by means of four primary strategies:

(1) analogy: describing similarities between familiar notions and new ideas;
(2) abstraction: introducing students to specific examples before making general claims;
(3) isomorphization: presenting a familiar concept or structure that is isomorphic to the new one at hand[1];
(4) postponing: stating that the significance of a topic will be realized later when it is not currently obvious.

Indeed, matrix multiplication is undoubtedly an introductory topic in a first-semester linear algebra course (regardless of whether each textbook explicitly characterizes it as such). To this end, these four techniques provide an effective means with which to classify the rationale and explanations given for matrix multiplication.

Method

We narrowed our focus to introductory linear algebra textbooks (as advanced books are less likely to explicitly detail matrix multiplication) that had been published within the past decade (as these books are more likely to be in use in undergraduate classrooms). We compiled an initial list of recently-published textbooks by (1) examining syllabi available online for introductory linear algebra courses at more than 20 large universities around the United States, (2) conducting online searches of textbook provider websites, and (3) examining the textbooks in our own respective university libraries. Overall, our list is comprised of 17 modern, introductory linear algebra textbooks[2].

For each textbook, we examined any sections involving matrix arithmetic or matrix-vector products and also scanned the table of contents and index for any mention of these topics. Relevant pages were photocopied (or, for online books, printed out). Once all data had been

---

[1] Note that analogy and isomorphization seem quite similar. We shall distinguish the two by reserving isomorphization for literal cases of mathematical isomorphism; analogy is reserved for all other comparisons.
[2] Due to their propensity for introducing topics in very similar (if not identical) ways, textbooks sharing an author were deemed equivalent (and only counted once).

collected in this manner, each textbook was analyzed using the framework detailed above. The framework then enabled us to identify trends and common themes across the entire data set.

Results

Two primary methods of defining matrix multiplication were found[3]: (1) the *linear combination of the columns method* (LC) (in which the matrix vector product Ax is defined as a linear combination of the columns of A), and (2) the vector *dot product method* (DP).

Sequencing of content. Two primary sequences for developing matrix multiplication emerged. First, systems of equations were reframed as the matrix equation Ax=b, wherein Ax was defined as a linear combination of the columns of A. The more general matrix product AB was then defined in terms of the matrix-vector product. In this trajectory, the dot product method was secondary and arose as a means to calculate quickly or calculate a single entry (as opposed to the entire product matrix). Second, while systems of equations preceded matrix multiplication in each textbook we examined, some textbooks made no explicit mention of the relationship between their proposed definition of matrix multiplication and the systems of equations. These textbooks primarily presented matrix multiplication starting with the dot product method (on either the matrix-vector product or the product of two matrices). The linear combination of the columns method was given considerably less focus. We refer to these respective trajectories as (1) LC to DP and (2) DP to LC. The following table classifies each analyzed textbook[4]:

Sequence	Textbooks Employing Specified Sequence
LC to DP (7 total)	Cheney & Kinkaid (2012); Holt (2012); Lay (2011); Leon (2010); Nicholson (2013); Spence, Insel, & Friedberg (2007); Strang (2009)
DP to LC (10 total)	Andrilli & Hecker (2009); Anton & Rorres (2010); Bretscher (2012); DeFranza & Gagliardi (2008); Kolman & Hill (2007); Larson (2012)[5]; Poole (2011); Shifrin & Adams (2010); Venit, Bishop, & Brown (2013); Williams (2012)

Rationale. There were examples of each of the four categories of rationale. Typically, the rationale for the LC to DP method only included isomorphization (as solving the matrix equation Ax=b is isomorphic to solving the system itself). In contrast (and somewhat interestingly), the DP to LC method spanned the remaining three categories.

Rationale	Method	Example
Analogy	DP to LC	"Another basic matrix operation is matrix multiplication. To see the usefulness of this operation, consider the following application in which matrices are helpful for organizing information. A football stadium has three concession areas …" (Larson, 2012, p. 42)
Abstraction	DP to LC	"Another useful operation is matrix multiplication, which is a generalization of the dot product of vectors." (Andrilli, 2009, p. 59)
Isomorphization	LC to DP	"Theorem 3 provides a powerful tool for gaining insight into

[3] There are, of course, other methods that can be used to multiply two matrices. Those listed are the most prevalent among the textbooks we examined. For additional information about the nonstandard methods, see Carlson (1993) and Strang (2009). A common method usually occurring amongst the more advanced material in these texts is to link matrix multiplication to preserving the composition of linear transformations.
[4] While each textbook has its own style and slight variations from the other texts, this table displays the general, overall pattern of each.
[5] Larson, Edwards, & Falvo (2008) proceeded identically.

		problems in linear algebra, because a system of linear equations may now be viewed in three different but equivalent ways: as a matrix equation, as a vector equation, or as a system of linear equations." (Lay, 2011, p. 36)
Postponing	DP to LC	"Only a thorough understanding of the composition of functions and the relationship that exists between matrices and what are called linear transformations would show that the definition of multiplication given previously is a natural one." (Kolman & Hill, 2007, p. 24)

Discussion and Questions

While analysis of the data is still ongoing, there is sufficient evidence here to conclude that these introductory textbooks, while presumably having similar goals, enact strategies with contrasting objectives and rationale to achieve them. We hope to engage the audience in a discussion about the pedagogical implications of these differing content sequences and which approach is more effective. In particular, we will ask our audience following questions:

- Which method of defining matrix multiplication do you prefer? Why?
- Does one of the two sequences make more sense mathematically? Why?
- While the effectiveness of the two approaches has not been evaluated (quantitatively or otherwise), can a sound argument be made that one is more effective than the other?

References

Andrilli, S. & Hecker, D. (2010). *Elementary linear algebra* (4th ed.). Massachusetts: Academic Press.

Anton, H. & Rorres, C. (2010). *Elementary linear algebra: Applications version* (10th ed.). Hoboken, NJ: Wiley.

Bretscher, O. (2012). *Linear algebra with applications* (5th ed.). New Jersey: Pearson.

Carlson, D. (1993). Teaching linear algebra: Must the fog always roll in? *The College Mathematics Journal, 24(1)*, 29-40.

Carlson, D., Johnson, C.R., Lay, D.C., & Porter, A.D. (1993). The linear algebra curriculum study group recommendations for the first course in linear algebra. *The College Mathematics Journal, 24(1)*, 41-46.

Cheney, W. & Kinkaid, D. (2012). *Linear algebra: Theory and applications* (2nd ed.). Massachusetts: Jones & Bartlett Learning.

DeFranza, J. & Gagliardi, D. (2008). *Introduction to linear algebra*. New York: McGraw-Hill.

Kolman, B. & Hill, D. (2007). *Introductory linear algebra* (9th ed.). New Jersey: Pearson.

Harel, G. (1987). Variations in linear algebra content presentation. *For the Learning of Mathematics, 7*, 29-32.

Holt, J. (2012). *Linear algebra with applications*. New York: W.H. Freeman.

Larson, R. (2012). *Elementary linear algebra* (7th ed.). Massachusetts: Houghton Mifflin.

Larson, R., Edwards, B. & Falvo, D. (2008). *Elementary linear algebra* (6th ed.). Massachusetts: Houghton Mifflin.

Lay, D. (2011). *Linear algebra and its applications* (4th ed.). New Jersey: Pearson.

Leon, S. (2010). *Linear algebra with applications* (8th ed.). New Jersey: Pearson.

Nicholson, K. (2013). *Linear algebra with applications* (7th ed.). New York: McGraw Hill.

Poole, D. (2011). *Linear algebra: A modern introduction* (3rd ed.). Massachusetts: Houghton

Mifflin.
Spence, L., Insel, A. & Friedberg, S. (2007). *Elementary linear algebra: A matrix approach* (2nd ed.). New Jersey: Pearson.
Strang, G. (2009). *Introduction to linear algebra* (4th ed.). Massachusetts: Wellesley Cambridge Press.
Shifrin, T. & Adams, M. (2010). *Linear algebra: A geometric approach* (2nd ed.).. New York: W.H. Freeman.
Venit, S., Bishop, W., & Brown, J. (2013). *Elementary linear algebra* (2nd ed.). Ontario: Nelson Education.
Williams, G. (2012). *Linear algebra with applications* (8th ed.). Massachusetts: Jones & Bartlett Learning

STUDENT UNDERSTANDING OF MEAN, DISTRIBUTION AND STANDARD DEVIATION

Samuel A. Cook — Wheelock College
Tim Fukawa-Connelly — Drexel University

This study investigates the understandings of mean, median, distribution and standard deviation that undergraduate students have at the end of an introductory statistics course. The goal was to explore their understandings as a follow-up to previous studies documenting incoming student difficulties with the concepts and determine whether a course would help them achieve a more statistically appropriate understanding. They overwhelmingly think about the mean as the "average" and via the calculating formula, meaning they understand it as a process. Similarly, they understand the median in terms of the process for determining it, or via the location-based term, "middle." As a result, students do not generally understand the two measures to be describing a similar concept. Students do, reliably connect the shape of a distribution to standard deviation, but that connection varies by type of display and is not based on a reliable rule.

Key words: measures of center, measures of variation, student understanding, statistics education

One of the most important, and difficult ideas for students to develop is the concept of a sampling distribution (Chance, delMas, & Garfield, 2004; delMas, Garfield, & Chance, 2004). Research suggests students' difficulties with the foundational concepts of center, distribution and variation is the primary contributor (delMas & Liu, 2005).

The Arithmetic Mean

The arithmetic mean is a fundamental topic in terms of measures of center in the development of statistical knowledge and sophistication (Landrum, 2005; Mokros & Russell, 1995). The mean can be understood in a variety of ways. Consequently, it would be beneficial for students to have a variety of possible conceptualizations of the mean (Cook & Fukawa-Connelly, 2012; Watier, Lamontagne, & Chartier, 2011). There is evidence that K-12 students often develop a very procedural understanding of the mean rather than one that includes an intuitive or conceptual understanding (Watier, Lamontagne, & Chartier, 2011) or one that is able to be operated on as an object (Mokros & Russell, 1995). At the conclusion of Mokros and Russell's study (1995) none of the students in the study showed a thorough understanding of the mathematical relationships of the mean and stated "One of the questions that remains is how children can ultimately come to understand the mean as a mathematical point of balance (p. 38)." Mokros and Russell therefore claim that the arithmetic mean, as an object, should not be introduced until after students have mastered other fundamental ideas. The inability for 8th graders to arrive at an abstract understanding of the mean is unsurprising when considered with what is known about cognitive development (Woolfolk & Perry, 2012). In particular, there is significant evidence that high school students, and even many adults, do not have the ability to use formal synthetic reasoning (Shayer, 2003). It is much more typical that the type of formal operational ability needed to support an object-understanding of the mean can only come as a result of undergraduate instruction (Woolfolk & Perry, 2012). By the end of an introductory statistics course, students should have the reasoning ability and be developing abstract notions of the mean and be able to operate upon it. One of the goals of this study is to test that claim.

The Standard Deviation

Compared to students' understanding of measures of center, and the mean in particular, relatively little research has been conducted on students' understanding of variation and distribution (delMas & Liu, 2005; Peters, 2011). Yet, as Peters, drawing on delMas and Liu's work, explains, "understanding standard deviation or mean absolute deviation necessitates a dynamic conception of distribution that coordinates changes to the relative density of values about the mean with their deviation from the mean" (p. 55, 2011). That is, understanding the standard deviation both requires an ability to think about the mean in quite sophisticated ways and also about distribution in dynamic ways, both of which are believed to be difficult. There are two relevant studies to student understanding of the standard deviation. The first focused on students' initial understandings of standard deviation, prior to any undergraduate coursework and described stages of understanding that students might develop based upon asking them to complete a series of visualization tasks (delMas & Liu, 2005). Peters' paper was also describing different ways that students might reason about the standard deviation. Both were as focused on the ways that their participants could develop their thinking as describing how students are most likely to think.

Research questions

We respond to delMas and Liu's call for more study of how students " make comparisons of variation between two or more distributions" (p. 56). We intend to focus on the first piece, attempting to describe the ways of understanding measures of variation that the broad range of students hold while also expanding on extant studies of student understanding of the mean and measures of center. Therefore, we investigate the following research questions:

1. In what ways do undergraduates, who have completed an introductory statistics course, think about the mean and median?
2. Are undergraduates who have (nearly) completed an introductory statistics course able to determine which graphical representation shows a distribution with more variability

Previous statistics research has shown the utility of Sfard's (1991) description of two ways to understand a mathematical concept; as a process and as an object. In holding a process conception, an individual sees a concept as a "potential rather than actual entity, which comes into existence upon request in a sequence of actions" (p. 4). The second way that an individual might see a mathematical entity is as an *object*, referring to it as if it were a real thing that exists somewhere. Holding an object conception also means being able to apprehend the entirety of the concept, manipulate it as a whole, and operate or perform processes on the concept. The student can also explore the general properties of the category of concepts and relations between the concepts themselves. In general, holding a misconception can be understood as a normal part of the learning process, and development of more correct concepts may involve students concurrently holding two or more competing conceptions (Smith, diSessa, & Roschelle, 1994). This research should be understood as an investigation of the types of understandings of mean and standard deviation that students hold at the end of an introductory course in statistics, or, the most mature understandings that most students will ever hold. We investigate whether they have process or object conceptions, or, if they have gone down a side-route and have fixed on a quasi-structural set of ideas.

Methods

Data for this study was drawn from 41 participants from multiple universities. Each of the participants completed a 5-item survey that included three items focused on students' understanding of the mean and median and two items focused on students' understandings of standard deviation. The research team drew on items focused on student understanding of the mean and median that had previously been used in similar studies (Kim, Fukawa-Connelly, & Cook, 2012). The two items assessing student thinking about distribution, and the standard deviation, were taken from delMas and Liu's (2005) work. These two items were selected because they were best at differentiating students in terms of their thinking about distribution and standard deviation. Moreover, the items also consistently solicited different types of thinking about the concepts, thereby allowing the researchers to give more complex descriptions of student understanding.

Approximately two thirds of the students completed the survey on-paper while the rest of the participants completed an online version that had the same format except that all items were presented on a single page rather than front-to-back. The online survey did have a relatively large number of individuals agree to participate and then not complete any items. We believe a reasonable fraction of these were likely the individuals to whom we sent the original solicitation for recruitment of participants; departmental administrators and faculty, although we cannot be sure.

Survey responses were summarized in two ways. The first was used on survey items 1 and 2, students were asked to list all of ways of they think about the mean (item 1) and median (item 2). Responses were divided into six response categories for item 1 and seven for item 2. If a participant listed a way of thinking related to a category they were placed in that category, such that, each participant was placed in at least 1 category, but may have been placed in every category if their list dictated such. The second was used on item 3 where participants were asked if the mean and median describe similar things and items 4 and 5 where participants were asked to identify the larger standard deviation between two visual summaries. Each response was placed into a single category response that identified both the answer the participant chose and the reasoning they gave. From these categorizations, we identified individual responses that were representative and interpreted them in terms of the categorization appropriate for either mean (Cook & Fukawa-Connelly, 2012) or standard deviation (delMas & Liu, 2005) and explain the type of understanding that the students most likely hold.

Data and Results

Of the 41 participants, 39 chose to provide a response to item 1, which asked participants to "Please list all the ways that you know how to describe or think about the mean"; these responses are summarized in table 1. The most common response participants gave in item 1 was to describe the mean as an average as 35 of the 39 participants used average, in some way, in their response. Another common response was to include the formula for the arithmetic average, either symbolically, through example or through an explanation such as "add up the numbers and divide the amount"; 17 of the 39 respondents thought of a mean as a formula, with 16 of the formulas deemed correct and one that stated that he/she thinks of a mean as "add up and average of all data". Of the 17 participants that think of a mean as a formula, 15 of them also stated that they think of a mean as an average, and 13 of the 17 participants did not have any additional ways of thinking about the mean. Of the 39 participants 23 (59%) think about the mean as only an average, only a formula or an average and a formula, but not more.

Table 1. Responses about the mean

Response	Average	Formula	Center, Middle or Representativness	Samples vs Populations	Arithmetic vs Geometric	Other
Percentage of Participants	89.7%	43.6%	28.2%	10.3%	2.6%	15.4%

Of the 41 participants, 39 provided a response to item 2, which asked participants to "Please list all the ways that you know how to describe or think about the median"; these responses are summarized in table 2. In some way, 82.1% of responses included a reference to the middle of data with 53.8% of respondents being more specific by indicating an ordered middle, explaining the procedure to find the median or doing both. In addition to the 82.1% who used the word middle, 10.3% did not use the word middle in how they think about the median, but did describe the median using the words center, central tendency or 50th percentile. There were 3 (7.7%) participants that did not provide a response that correctly described the median, 2 described the mode and 1 stated it was "the average number out of a few different statistical numbers". There were also 3 participants who thought of the median as an "average", where average was used in a colloquial way, for example: "Median is the "middle" number of the data set. It also can reflect the average standard of the data.".

Table 2. Responses about the median

Response	Middle	Ordered Middle	Formula or Procedure	Center or Central Tendancy	Resistent to outliers	"Average"	Something incorrect
Percentage of Participants	82.1%	38.5%	30.8%	15.4%	10.3%	7.7%	12.8%

Of the 41 participants, all 41 chose to provide a response to item 3, which asked participants "Do you believe that the mean and median describe similar things?"; these responses are summarized in tables 3a and 3b. In the online version of the survey, participants were forced to pick between "yes" and "no" if they chose to answer the question. The paper surveys asked for a yes or no, but some participants said "sometimes yes and sometimes no"; therefore, if in the reasoning given on the online surveys it was indicated that the answer is sometimes yes or sometimes no, they were placed in that answer category.

When participants responded that the mean and median sometimes do and sometimes do not describe similar things, in all cases their reasoning was based on the result of the calculation or the final "figure" as several students wrote. Some participants went further and indicated that the figures would only be similar if the data was normal or symmetric.

Focusing on the result or final "figure" was also popular reasoning for answering both yes and no. For those who indicated that yes, they do describe similar things, 16.7% noted that the result of the mean and median are usually close together; therefore, they are similar. Likewise, of those who indicated that they did not believe they describe similar things 46.7% reasoned that because the results will not always be close together they are not describing similar things; some added further reasoning around outliers or issues of skewed distributions.

Among all respondents, 43.9% of them used the result of the calculation of the mean or median as the reason for selecting yes, no, or both yes and no. Another reason students claimed that the mean and median do not describe similar things drew on the definitions of the two, one is an average and the other a middle point and are therefore not describing similar things.

The reasoning for participants determining that the mean and median describe similar things is summarized in table 3b. The most detailed reasoning was found for those who responded with a yes, with 55.5% of yes respondents citing that both describe the middle or center of data, or the notion of representativeness was used. One response indicated that both represent an "average", where average represented the colloquial definition and not the mathematical or statistical. Additionally, 16.7% of respondents did not provide a clear reason but did specifically indicate that they recognized a mean and median are not the same thing, but do describe similar things.

Table 3a - "Do you believe that the mean and median describe similar things?"

Response	Yes	No	Sometimes yes and sometimes no.
Percentage of Participants	43.9%	36.6%	19.5%

Table 3b - "The mean and median describe similar things"

Reasoning	Center, Middle or type of "Average"	Numerical value is similar.	Similar but not Same, with no additional reasoning.	No reasoning or incorrect reasoning
Percentage of Participants	55.5%	16.7%	16.7%	11.1%

Of the 41 participants, all 41 provided a response to item 4, which asked participants "When determining the size of the variability of the data summarized by the following two histograms, is the standard deviation for the graph on the left smaller, larger or the same as the graph on the right?"; there were 2 responses that did not provide a clear answer between smaller, larger or the same so those were ignored and a sample of 39 is summarized in table 4a. The histograms described two distributions of similar symmetric shape and mode, but with differing areas of no data such that the graph on the left had a smaller standard deviation. Of participants who concluded the histogram on the left had a smaller standard deviation (69.2%), 55.6% of them reasoned similar to delMas and Liu's study, giving a reason that was categorized as there being more data near the middle of the left histogram than the right; while 11.1% used a mathematical reason similar to "more in the middle", reasoning that the deviations on the left would have more small differences. These results are summarized in table 4b.

Of those respondents who did not chose smaller, the majority decided that the standard deviations were the same with the primary reason given that both histograms were balanced or, similarly, that the mean and median were the same. These participants concluded that the actual deviations on the right and on the left were identical and thus the variation of each was the same. These results are summarized in table 4c.

Table 4a - "The standard deviation for the above graph on the left is:"

Response	Smaller than the graph on the right.	Larger than the graph on the right.	the same as the graph on the right
Percentage of Participants	69.2%	5.1%	25.6%

Table 4b - "The standard deviation for the graph on the left is smaller than the graph on the right"

Reasoning	More in the Middle	Use of formula or mathematical reasoning.	Different distribution or left has a steeper peak in distribution	No reasoning
Percentage of Participants	55.6%	11.1%	11.1%	22.2%

Table 4c - "The standard deviation for the graph on the left is the same as the graph on the right"

Reasoning	Mean = Median	Balanced Distribution or distance from mean the same on both sides	Follow the same distribution shape.	No reasoning
Percentage of Participants	30%	40%	20%	10%

Of the 41 participants, 38 provided a response to item 5, which asked participants "When looking at the box-plots above, is the standard deviation for the data in the graph on the left smaller, larger or the same as that in the graph on the right"; in 3 of the responses an answer could not be determined and are ignored, these results of 35 responses with determinable answers are summarized in table 5a. The box-plots described two distributions of similar medians but differing ranges and locations; such that, the box plot on the left had a smaller standard deviation. Of the participants who stated that the graph on the left has a smaller standard deviation 71.4% of them referenced the difference in range, IQR, spread or some combination of range, IQR and spread as their primary reasoning, and 14.3% referenced more data near the middle. Other than participants that did not provide a reason, this accounted for everybody who chose the box-plot on the left.

There were 7 remaining participants who did not select the box-plot on the left as smaller, 5 of them provided no reasoning or reasoning that was difficult to follow, 1 indicated a difference in the size of the IQR, but then concluded the left was larger, and one used differences in skewedness to indicate the box-plot on the left had a larger standard deviation.

Table 5a - "The standard deviation for the above graph on the left is:"

Response	Smaller than the graph on the right.	Larger than the graph on the right.	the same as the graph on the right
Percentage of Participants	80%	14.3%	5.7%

Table 5b - "The standard deviation for the graph on the left is smaller than the graph on the right"

Reasoning	Smaller IQR/Range/Spread	More in the middle, no reference to range or spread.	No reasoning
Percentage of Participants	71.4%	14.3%	14.3%

Discussion

The results presented above demonstrate how students think about the mean, median, and variation. In terms of the mean, the results show that the students overwhelmingly think about the mean via the term average and via the formula. Cai, Lo and Watanabe (2002) showed that textbooks in the United States present the formula (a process-based understanding), which is contrasted with an equal-sharing (process or object) or per-unit quantity type of understanding (object) and at a more advanced level, the mean can also be understood as a representative of the data set (object). That the students typically hold a process-conception of the mean was reinforced by the results about the median (almost always described in terms of "location" or a process) in that the students' discussion about whether the mean and median could be said to describe the same thing focused on numerical result, only claiming that they would be similar if the results are similar. That is, because the processes yield different results, the concepts do not describe similar things. This finding demonstrates that a semester of undergraduate statistics was insufficient to help students progress to more abstract understandings, even though they are needed in order to develop correct notions of inferential statistics.

In terms of the students' understanding of variation and distribution, this study was designed to determine whether students at the end of a traditional undergraduate statistics course had understandings of spread and variation more similar to the pre-intervention or post-intervention students in delMas and Liu's (2005) study. In general, the students in our study did primarily draw upon reasoning related to spread or variation. Overall, this is the type of reasoning that we want to promote (even if the results of that reasoning were incorrect for many of them), and, it suggests that the students recognize that the shape of the distribution does give information about the standard deviation. It is important to help students better reason in that way, meaning, to correctly recognize spread and variation from a histogram and a box-and-whisker plot. This study has the limitation of not being able to further explore or explain the students' reasoning, and, as a result, we do not attempt further inferences about it.

There has been, thus far, no direct investigation attempting to document classroom instruction in statistics and link it to changes in student understanding (there have been multiple teaching experiments). Our work suggests a need to attempt to do so, perhaps as a means to motivate adoption of research-based curricula, but at the least, in order to explain why students hold the conceptions that they do.

References

Cai, J., Lo, J. J., & Watanbe, T. (2002). Intended treatments of arithmetic average in U.S. and Asian school mathematics textbooks. *School Science and Mathematics, 102*(8), 391-406

Chance, B., delMas, R. C., & Garfield, J. (2004). Reasoning about samling distributions. In D. B.-Z. J. Garfield (Ed.), *The challenge of developing statistical literacy*: The Netherlands: Kluwer Academic publishers.

Cook, S. & Fukawa-Connelly, T. (2012). *Toward a Theory of Symbol Sense in Undergraduate Statistics.* Proceedings of the SIGMAA RUME*: Conference on Research in Undergraduate Mathematics Education*. Portland, OR. http://sigmaa.maa.org/rume/crume2012

delMas, R. C., Chance, B. L., & Garfield, J. (1999). A model of classroom research in action: Developing simulation activities to improve students statistical reasoning. *Journal of statistics education, 7*(3).

delMas, R. & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55-82.

Garfield, J. & Ahlgren, A. (1988). Difficulties in learning basic concepts in statistics: Implications for research. *Journal for Research in Mathematics Education. 19*, 4443.

Garfield, J. & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice.* New York, NY: Springer.

Kim, H., Fukawa-Connelly, T., & Cook, S. (2012). *Student Understanding of Symbols in Introductory Statistics Courses.* Conference on Research in Undergraduate Mathematics Education. Portland, OR.

Mokros, J. & Russell, S. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, 26(1), 20-39.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, *65*, 123-165.

Peters, S. (2011). Robust understanding of statistical variation. *Statistics Education Research Journal*, *10*(1), 52-88.

Shaughnessy, J. M. (1985). Problem-solving derailers: The influence of misconceptions on problem-solving performance. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving* (pp. 399-415). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shayer, M. (2003). Not just Piaget; not just Vygotsky, and certainly not Vygotsky as alternative to Piaget. *Learning and Instruction, 13*, 465-485.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

Smith, J.P., diSessa, A., & Roschelle, J. (1994). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *Journal of the Learning Sciences*, *3*(2), 115-163.

Watier, N., Lamontagne, C., & Chartier, S. (2011). What does the mean mean? *Journal of Statistics Education*, *19*(2). Retrieved from http://www.amstat.org/publications/jse/v19n2/watier.pdf

Woolfolk, A. & Perry, N. (2012). *Child and adolescent development*. Upper Saddle River, NJ: Pearson.

107

A TYPOLOGY OF VALIDATING ACTIVITY IN MATHEMATICAL MODELING

Jennifer A. Czocher
Texas State University

Mathematical modeling tasks are used to help students learn mathematics and also to improve their modeling skills. Validating *has been identified as the process by which students check and revise their models, but little is known about when or how students choose to do so. This study examined engineering students' validating activity and characterized it into a typology of different kinds of validating activity satisfying different roles in ensuring accuracy of the model.*

Key words: differential equations, mathematical modeling, qualitative methods

Introduction

Mathematical modeling tasks have been shown to help students learn significant mathematical ideas (Lesh, Cramer, Doerr, Post, & Zawojewski, 2003) and to hone students' modeling skills (Haines, Crouch, & Davis, 2001; Mousoulides, Christou, & Sriraman, 2008). Researchers have described students' modeling activity an iterative process, where the connection between the real world and the mathematical world is continually refined (Blum & Leiß, 2007). While theory emphasizes the importance of the iterative refinement of the mathematical representation and the modeler's interpretation of the situation (Lesh & Yoon, 2007; Thompson & Yoon, 2007), it has not yet clarified how such refinements or reinterpretations occur. Current descriptions of how modelers revise their models rely on the activity *validating* – confirming that a model yields accurate predictions – but there is not yet a description of how students validate their models. The present study addressed this gap.

The research presented here is part of a larger study of engineering students' mathematical thinking during mathematical modeling, but the current focus is on characterizing students' validating activity. Drawing on evidence generated by undergraduate engineering majors working on mathematical modeling tasks, this report elaborates on validating activity and addresses the research question: *How do engineering majors' validating activities contribute to the refinements of their mathematical models*?

Theoretical Framework

The construct *validating* is drawn from Blum & Leiß's (2007) theoretical model of individuals' mathematical modeling activity, which served as the research framework for this study. A schematic of the research framework is shown in Figure 1. The framework separates the world into the real world and the mathematical world and overlays an iterative cycle of activity (separated into stages of model building [a – f] marked in red and transitions among those stages [1-6] marked in blue). The cycle represents the transformation of a real world problem into a mathematically well-posed problem (stages [a – d], transitions [1 – 3]), which is analyzed mathematically (stages [d – e], transition [4]), interpreted in terms of the real world (stages [e-f], transition [5]), and then compared against the state of the real world (transition [6]).

According to the framework, *validating* consists of verifying results obtained from analysis and interpretation of a mathematical model against measurements from the real world (transition [6]). This definition can be broadened by recognizing that it is essentially a comparison activity: comparing an obtained outcome (object of validation) to an expected outcome (standard of validation). The result is that the model is either accepted or rejected. When the results are accepted, the modeler exits the modeling cycle. If the model is rejected, the theoretical framework predicts that the modeler re-enters the modeling cycle after making adjustments to the model.

Methods

The stages and transitions outlined by the theoretical framework guided task design, data analysis, and data interpretation. Mathematical modeling tasks were designed to elicit each of the six transitions predicted by the framework. Seven one-on-one, semi-structured, modeling task-based interviews were conducted with four undergraduate engineering majors (total of 60 hours of interviews) who were purposefully selected according to their performance on a knowledge of calculus screening test. Mance and Trystane were selected for low performance while Torrhen and Orys were selected for high performance. They were in their freshman and sophomore years and enrolled in a differential equations course at a large Midwestern university. Each session lasted between 1 and 2 hours and addressed between 1 and 3 tasks. Mathematics elicited by the interview modeling tasks ranged from arithmetic to differential equations. The interviewer interacted with the students to request clarification of the students' solutions or to challenge the students' solutions. The sessions were audio/video recorded.

Data analysis followed a four stage process. First, indicators were developed from transcribed interviews and literature to identify where the students' utterances and mathematical work corresponded to transitions in the modeling cycle (samples given in Table 2). Second, using the indicators, students' utterances and mathematical work were tagged with the transitions in the modeling cycle. During data analysis it was observed that students' utterances and mathematical work corresponding to the middle of the modeling cycle also fit the descriptors for validating activity. Instances of validating activity were selected for systematic inspection. From this closer inspection, five kinds of validating activity were observed based on what stage of the mathematical model the student was validating (object of validation) and the stage of the mathematical model the student was comparing it to (standard of validation). See Table 3 for samples. Third, a typology based on the object/standard of validation was developed and used to characterize instances of validating activity. Fourth, second-order models of the students' mathematics were constructed (Steffe, 2013). These are essentially the researcher's narrative accounts of the students' mathematics as interpreted through the researcher's own mathematical theories and experiences. The narratives were used to ground the students' mathematics use in context and are what the illustrations presented here are drawn from. These second-order models emphasize the student's use of mathematical ideas rather than whether each step in solving the problem was correct. Therefore, during data analysis, the research attended to the students' validating activity and their own interpretations of the correctness of their own mathematics and modeling activity.

An example of the tasks used is the falling body problem, which is amenable to solution via kinematics, energy equations, and differential equations: *On November 20, 2011, Willie Harris, 42, a man living on the west side of Austin, TX died from injuries sustained after jumping from a second floor window to escape a fire at his home. What was his impact speed?*

Results

The data suggested that students used validating activity for two different, but related, purposes: to check the model's representativeness and to check the model's accuracy. Based on evidence from the interview sessions, the definition of *validating* activity was expanded into a typology of five types of validating activity (Table 3) which are illustrated below. An example is offered for each kind, but the kinds are not mutually exclusive. Due to space constraints, the examples were selected purposefully for their illustrative power so as to demonstrate each type as clearly as possible. Thus, the selection is not representative of all the tasks.

However, at least one example was chosen to share from each student and the illustrations are typical of the data set as a whole. The typology was capable of characterizing all instances of validating activity. Abridged task statements are given in Table 4. Throughout, italicized segments of protocol were coded as validating activity.

Type (i): Mathematical Results → Mathematical Model. In order to use a model to make accurate predictions, two conditions must be met. The first is that the mathematical model must correctly represent the situation. The second is that the analysis of the model that yields the prediction must be correct. Type (i) activity amounts to verifying the outcome of mathematical analysis and ensuring that the model satisfies the second condition.

In the Tropical Fish Tank problem, Mance set up a differential equation to model the amount of buffering agent in the tank. The model was incorrect, but he was able to solve it using integration techniques.

Mance: So $\frac{dc}{dt}$ umm, integrate that, you're gonna go with c equals, umm 5, no 60, no, plus now, negative t over 60. That's the integral. I'm pretty sure that's what you have to do for that. If this [the expression for $\frac{dc}{dt}$] is your, if that was your initial rate you were given, um, to get a concentration at a given point you need the concentration as a actual variable. *So I did the integral of this* [$\frac{dc}{dt}$] *with respect to t and then this will just be c.*

In this excerpt, Mance reviewed the product of his work, checking that he could see the connection between the starting expression and the resulting expression through his analysis.

Type (ii): Mathematical Model → Situation Model. Type (ii) activity is notable in that there is no mathematical analysis taking place. The modeler directly compares the mathematical model to his conceptual model of the problem setting. It contributes to ensuring that the model is representative of the real situation. In contrast, Type (iii) has to do with ensuring that relationships among variables in the mathematical representation reflect those chosen as important. In the Falling Body problem, Trystane compared the variables in his mathematical model to the quantities he expected to appear in it and found that his idealized real model was missing a variable. Using dimensional analysis to derive a differential equation, he noted that variables were missing from his model.

Trystane: Looks to me like we need the mass of something times a coefficient of the air – dunno why I made that a *W*.

Interviewer: You might've been calling it "wind."

Trystane: Oh, wind! The coefficient of the friction of the wind times velocity equals force. And this was force. *Intuitively, I don't think I trust that. That's* [the differential equation] *the answer that I reached. But I really think it has something to do with the surface area because this pencil would fall faster than a big piece of paper weighing the same amount.*

Here, Trystane directly compared his mathematical model to the situation model and found that it was lacking a variable that he expected would impact the force of drag experienced by the falling object. This prompted him to declare the variable – incorporate surface area into his idealized real model – and try to represent it in the mathematical representation for drag.

Type (iii): Mathematical Representation → Real Model. This type of validating is characterized by an effort to check whether the mathematical relationships among the factors present in the model match the intended physical relationships in the idealized real model. Type (iii) validates mathematizing activity and ensures that the model is accurate given assumptions made during simplifying/structuring activity. It tended to surface in tasks involving dynamics. In the Empire State Building problem, Orys decided that fitness of the individual ascending the building by stair must be accounted for. He invented a constant (F) to represent fitness and manipulated it and his model such that a less fit person would have a slower ascension pace.

Orys: Maybe just rearrange this a little. It's gonna be like steps time F over speed. I guess it would even be just put F on the bottom like $F \times speed$, based on the more fit you are the faster you're going. This is just speed, is the average speed, someone would travel and say you're more fit than the average person, you should go maybe 1.2 times that speed and you're less fit and you would go .8 times that speed.

Orys ensured that the principle *more fit people go faster* was captured by the product $F \times speed$. He did this by comparing his mathematical representation to the principles included in his idealized real model of the situation.

Type (iv): Real Results → Situation Model. This is the kind of validating activity predicted by theory. It compares the model's prediction against the modeler's expectations of the real situation and is used to ensure accuracy of the model. In the data, it tended to happen only when the task requested a numerical prediction from the derived model. In the Falling Body problem, Torrhen used kinematics equations to predict that a falling body would impact at 18.12 miles per hour. He validated this against his empirically based perceptions of speed:

Torrhen: If I were to hit a brick wall with that, it would hurt a lot or possibly kill me and 18 mph definitely. And it's [the prediction] higher than that. A car crash at 18 miles per hour isn't too much. But if you were completely unprotected, if you were hit by a car going 20 mph, not so much a car, but a bus that was more flat like going into the ground, then that would be painful.

Torrhen justified his obtained value by comparing it to a value he knew would cause the kind of damage to the falling body that was reported in the problem statement.

Type (v): Real Results → Real Model. This kind of activity was observed only in tasks that required differential equations but *did not* request a numerical prediction. This is because the result of analyzing a differential equation is a family of equations, in essence, another model. Trystane derived a first-order linear equation to model a falling body influenced by gravity and air resistance. As a solution, he obtained $Q(t) = Ce^{\lambda t}$. Unable to determine from the expression whether the exponential modeled the velocity of the falling body, he drew two graph shapes, one linear and one exponential. His dilemma was which type of relationship would best describe how force-due-to-drag would increase with velocity. On its own, this behavior demonstrates Type (iii) validating activity. However, it is embedded in a vignette of Type (v) activity. The exponential expression is the result of mathematical analysis and in the following excerpt, Trystane contemplated whether it fit his expectations from the real world.

The transcript below shows Trystane comparing his model to his expectations of how the important variables (force-due-to-drag and velocity of the object) should be related.

Trystane: *I'm not sure that's* [the exponential solution] *right because I'm not sure if there should be some sort of constant increase as you get faster*, um. I guess that just stems from fluid mechanics. For instance, I don't know if it's a linear graph [draws linear graph] or if as you're going faster it gets [draws exponential graph].

Interviewer: What's this graph represent?

Trystane: This [linear graph] is a straight line.

Interviewer: Of what versus what?

Trystane: Of force due to wind [vertical axis] and this is velocity [horizontal axis] and then this is force. I guess that would be switched [switches axis labels on exponential graph]. Force gets greater and greater.

Interviewer: So you're saying there's an increasing relationship between these two, force and velocity, but you don't know—

Trystane: -- I don't know exactly what this would look like and then this would determine whether this [the differential equation] equals zero or this equals some sort of forcing time. So I mean, right now I'm getting λ equals the negative of a constant. Um, but this doesn't really, it doesn't really tell me a whole lot because I don't know what the graph should look like. *I feel like it probably equals some sort of forcing term. Because I don't think the solution would end up being that as v* [velocity] *increases and position, I don't think it's gonna be $Ce^{\lambda t}$.*

In this excerpt, Trystane relied on graphical models of a qualitatively distinct relationship between the two variables of interest to make a decision about whether the exponential representation was acceptable. In essence, he compared his mathematical model with a condition of his real model. He decided against the exponential because he was concerned with how an exponential velocity would affect position.

Discussion and Conclusions

The typology described here is a product of a novel use of the mathematical modeling cycle developed by Blum & Leiß (2007). It was capable of characterizing all instances of validating activity observed during the interviews. The mathematical modeling cycle was built with the idea that validating occurs when checking the real results (stage [f]) against empirically collected data. Analysis suggests that this is not the only kind of validating activity and that other kinds of validating activities lead to model refinement and may even offer insight into the problem setting. In terms of the mathematical modeling cycle, the object of validation can be the situation model, the real model, the mathematical representation, the mathematical results, or the real results. The five kinds of validating activity could be separated into two interrelated purposes: those which examine predictions from the model (Types (i), (iv), and (v)) and those which examine fidelity of the model to the situation it was supposed to represent (Types (ii) and (iii)). Both are necessary for and contribute to developing accurate, representative mathematical models.

Longer and more involved tasks – those requiring substantial mathematical analysis – tended to prompt more comparisons between the mathematical result and the mathematical model (Type (iii)). In addition, tasks involving dynamics seemed to prompt more validation of the mathematical model itself against the real and situation models. Likewise, comparison of real results to the situation model – the kind of validating activity predicted by theory – happened when numerical predictions were made.

The current work may explain what are termed *blockages* in the modeling cycle (Galbraith & Stillman, 2006). Blockages were conceptualized as impediments to progress from one stage of the model to the next. The researchers presented a taxonomy of elements in the solution process that students could get caught up on and which would derail a solution procedure, such as "clarifying context of problem" (corresponding to real situation/situation model → real model transition) or "using correctly the rules of notational syntax" (corresponding to real model → mathematical results transition) (Galbraith & Stillman, 2006, p. 147). The researchers generated a modeling cycle that included backwards arrows at each site where blockages could occur, essentially reversing the transitions between stages in Blum & Leiß's (2007) model. Conceptualizing backwards motion through the modeling cycle as blockages is in part based on the assumption that validating can occur at only one site in the modeling cycle. The present analysis suggests that these blockages may be precipitated by validating activity. Future research should investigate whether and why an instance of validating activity may have led to a blockage. To do so, tasks which elicit different kinds of validating activity must be developed.

Another direction for future research would be to examine whether the individual's mathematical model is revised or if the individual's understanding of the problem situation is refined depending on the type of validating activity the student engaged in. Answers to such a question would aid in understanding the interplay between thinking about the real world and thinking about the mathematics that represents it.

This study has demonstrated that students engage in validating activity throughout the modeling cycle. Only checking the real results against the situation model (Type (iv)) is not sufficient for diagnosing *where* in the model-building process a mistake or misstep might have been made or how to rectify it. This is important for teaching with modeling because envisioning where students might need validation during a derivation or modeling problem can help in providing them support. This finding is important theoretically because the sites of validating activity may be sites where metacognition or coordination of different knowledge bases (e.g., scientific and mathematical) could be studied.

The present study suggests reivsions to the modeling cycle which would account for a variety of validating activities that are critical for refining mathematical models. Further research is needed to integrate the typology into the research framework and to understand which types of validating are most likely to occur and under what conditions.

References

Blum, W., & Leiß, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling (ICTMA 12): Education, Engineering and Economics* (pp. 222-231). Chichester: Horwood.

Galbraith, P., & Stillman, G. (2006). A Framework for Identifying Student Blockages during Transitions in the MOdelling Process. *Zentralblatt f\"{u}r Didaktik der Mathematik, 38*(2).

Haines, C., Crouch, R., & Davis, J. (2001). Understanding Students' Modelling Skills. In J. F. Matos, K. Houston, W. Blum, & S. P. Carreira (Eds.), *Modelling and Mathematics Education* (pp. 366-380). Chichester, England: Woodhead Publishing.

Lesh, R., & Yoon, C. (2007). What is Distinctive in (Our Views about) Models & Modelling Perspectives on Mathematics Problem Solving, Learning, and Teaching? In P. L. Galbraith, H. Hans-Wolfgang, M. Niss, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and Applications in Mathematics Education* (pp. 161-170). New York: Springer.

Lesh, R., Cramer, K., Doerr, H. M., Post, T., & Zawojewski, J. S. (2003). Model Development Sequences. In R. Lesh, & H. M. Doerr (Eds.), *Beyond Constructivism: Models and Modeling Perspectives on Mathematics Problem Solving, Learning, and Teaching* (pp. 35-58). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mousoulides, N. G., Christou, C., & Sriraman, B. (2008). A Modeling Perspective on the Teaching and Learning of Mathematical Problem Solving. *Mathematical Thinking and Learning, 10*, 293-304.

Steffe, L. P. (2013). Establishing Mathematics Education as an Academic Field. *Journal for Research in Mathematics Education, 44*, 354-370.

Thompson, M., & Yoon, C. (2007). Why Build a Mathematical Model: A Taxonomy of Situations That Create the Need for a Model to be Developed. In E. H. Richard A. Lesh (Ed.), *Foundations for the Future in Mathematics Education* (pp. 193-200). Mahwah, NJ: Routlege.

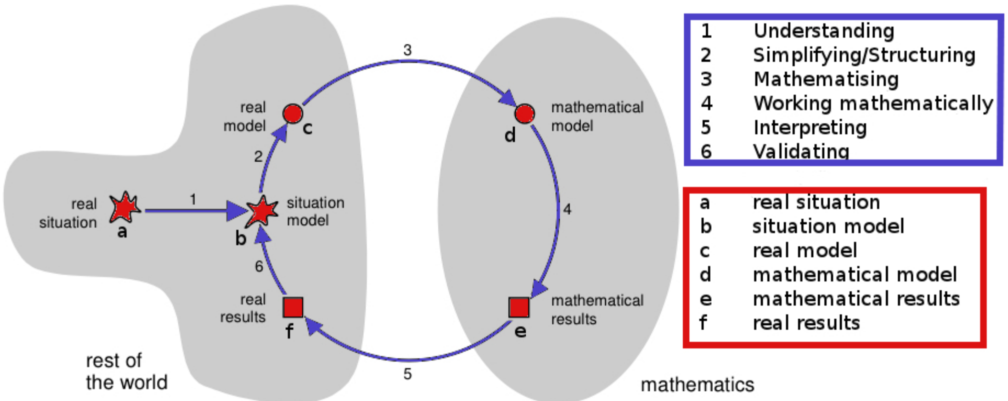


Figure 1: Research framework (Blum & Leiß, 2007)

Table 1 Stages of model building

Stage of Model	Definition
[a] real situation	situation, as observed in the world
[b] situation model	conceptual model of problem
[c] real model	idealized version of the problem (serves as basis for mathematization)
[d] mathematical model	model in mathematical terms
[e] mathematical results	answer to mathematical problem
[f] real results	answer to real problem

Table 2 Transitions and sample indicators

Transition	Captures	Sample Indicator
[1] understanding	forming an idea about what the problem is asking for	reading the task
[2] simplifying & structuring	identify critical components of the problem situation	making assumptions to "simplify" the problem
[3] mathematizing	represent the idealized real model mathematically	writing mathematical representations of ideas
[4] working mathematically	mathematical analysis	explicit algebraic or arithmetic manipulations
[5] interpreting	recontextualizing the mathematical result	speaking about results in context of the problem
[6] validating	verifying results against the real world	implicit or explicit statements about the reasonableness of the answer

Table 3 Typology of validating activity

	Type of Validating Activity		
	Object of Validation	Standard of Validation	Description
(i)	mathematical results	mathematical model	checking the results of a calculation or mathematical analysis mathematically
(ii)	mathematical model	situation model	comparing the mathematical model, its constituent components or relationships, to the interpretation of the problem setting
(iii)	mathematical model	real model	comparing the mathematical model, its constituent components or relationships, to the idealized version of the problem setting
(iv)	real results	situation model	the kind of validation predicted by the theoretical

			framework
(v)	real results	real model	comparing the real results against physical principles present and accounted-for in the real model

Table 4 Task the example was drawn from

Type of Validating	Task/Student	Abridged Problem Statement
(i)	Tropical Fish Tank/Mance	Model the amount of a pH buffering agent currently in a tropical fish tank.
(ii)	Falling Body Problem/Trystane	What is the impact velocity of a falling body?
(iii)	Empire State Building/Orys	How long does it take to ascend the Empire State Building via the tourist elevator? Via the stairs?
(iv)	Falling Body Problem/Torrhen	What is the impact velocity of a falling body?
(v)	Falling Body Problem/Trystane	What is the impact velocity of a falling body?

USING THE FLIPPED MODEL TO ADDRESS COGNITIVE OBSTACLES IN DIFFERENTIAL EQUATIONS

Jenna Tague
The Ohio State University

Jennifer Czocher
Texas State University

Greg Baker
The Ohio State University

Key words: Flipped classroom model, cognition, post-secondary education

Recent work has shown that there is a lack of coherence from calculus to differential equations (Czocher, Tague, & Baker, 2013). We define lack of coherence as the gap between the knowledge students are expected to gain by the end of the calculus sequence versus how calculus knowledge is expected to be used in differential equations (Czocher, Tague, & Baker, 2013). In this report, we describe how we have exploited the flipped classroom model to begin to address some of these issues with coherence. We share our theoretical perspective, how it was enacted, and also a preliminary evaluation of students' perceptions of the coherence of the course and its content.

Theoretical and Empirical Background

Two kinds of coherence arise in this paper. First, we address *coherence in the curriculum* through our planning and sequencing of mathematical content. Second, we address *coherence within instruction* by aligning out-of-class with in-class material and concepts. As such, our theoretical and empirical background speaks to both kinds of coherence.

Curricular Coherence

We draw on the construct *cognitive obstacle* (Herscovics, 1989) to describe a manner of thinking about a mathematical object or structure that is appropriate in one case, but inappropriate in another. Cognitive obstacles may arise from or contribute to incomplete concept images and may be a symptom of lack of coherence in the curriculum. Two well-documented cognitive obstacles relevant to differential equations learning are the *function-as-solution dilemma* (Rasmussen, 2001) and viewing *rate of change* as a symbolic process devoid of quantities (Zandieh, 2000; Rowland, 2006). These obstacles threaten to inhibit differential equations learning because they have to do with the nature of equations and the nature of the derivative. The function-as-solution dilemma refers to stuents' difficulty in thinking about solutions to differential equations because they are families of functions, rather than numbers (which solve algebraic equations). It is a cognitive obstacle because a limited understanding about solutions to differential equations can hinder development of ideas about equilibrium solutions and many others.

Student difficulty with rate of change as related to quantities is documented throughout mathematics learning, but it is a critical concept because of its relation to the derivative. To complicate matters, the derivative concept was also identified by (Czocher, Tague, & Baker, 2013) as being used differently in differential equations than it was expected to be known at the end of calculus due to its dependence on conceptions of rate of change of physical quantities. Other related mismatches have been reported between how engineering faculty and mathematics faculty teach and use the concept of derivative (Bingolbali, Monaghan, & Roper, 2007). Other lines of inquiry into student conceptualization of rate of change have revealed that they have difficulty isolating rate of change as a quantity of interest (Monk, 1992), a poor understanding of covariation and derivative (Thompson, 1994), and refer to derivative *as* the tangent line (rather than slope of the tangent line) despite defining derivative formally (Zandieh, 1997). With these issues in mind, we sought to use the flipped classroom model to strengthen students' conceptions of key calculus ideas known to be cognitive obstacles to learning differential equations content.

Since these two concepts are foundational to differential equations, we chose to focus our efforts on strengtheninng and broadening students' understanding of them. We did so by making them the focus of pre-class instructional modules. In-class instruction was adjusted to incorporate group work and discussion and the use of traditional lecture was minimized. In this paper, we focus on the pre-class modules and how it related to the content during the classtime rather than the pedagogy.

Instructional Coherence

During the 1990s, instructors began to use video, presentation software, and Internet capabilities to teach with the "inverted classroom" (or "flipped classroom") model. The philosophy was that students would go through lecture materials at home (e.g., watch a video) and class time would be used for what was traditionally at-home work (e.g., solving problems). In this way, students would be ready to engage with the most difficult parts of the content while the instructor was present. For most instructors, using a combination of video recordings, quizzes, and educational videos (e.g., TED talks) to move lecture out-of-class was not challenging. How to spend class time became problematic. In many flipped classroom settings, instructors lectured during the in-class time too. Other reported difficulties included: failure to address student misconceptions, overuse of low cognitive-level activities requiring only recall of facts, and an emerging disconnect between lecture materials and active-learning in-class components (Andrews, Leonard, Colgrove, & Kalinowski, 2011). In our use of the flipped classroom paradigm, we explicitly focused on addressing the difficulties reported by Andrews, et al (2011) for the purpose of exposing cognitive obstacles.

The scope of this paper is an examination of the coherence of our modified differential equations course from the students' points of view. We report a study of students' perceptions of the coherence of our implementation of the flipped classroom model and coherence of the resulting curriculum. We use this study as a starting place toward resolving coherence issues identifed in the literature – curriculuar coherence and course coherence.

Context

The differential equations course was a class of 80 second-year engineering majors. We identified target concepts (e.g. rate of change) through literature searches and observation of student work over the past four years. Modules were created using Articulate Storyline software which was embedded in the course website. Each module addressed an aspect of a cognitive obstacle for the upcoming differential equations content. Long- and short-answer questions were designed to encourage critical thinking about the cognitive obstacle or misconception rather than just working related procedural problems. Many of the modules had branches where question $n+1$ depended on the student's response to question n.

For example, the first module was designed to address one basic idea supporting *rate-of-change as a quantity*: what are some ways one can measure change in a quantity? The module asked students to examine data derived from a loan repayment schedule based on an exponential model using both absolute change and relative change. The students were first shown a table of values without context. The students were then asked to draw on their past experiences with measuring change to make sense of trends in the table. After asking the students to describe change in the data set in words, we introduced multiple representations of the data set including sequence and graphical representations. They were asked to consider how change in a quantity could be represented in representations other than algebraic. Interactions were shaped around helping the students articulate the idea that different representations highlight different attributes of the data

In class, the students were given the data shown in Figure 1. The table represented growth of bacteria in the same medium at different temperatures. The professor opened the course by asking, "What is happening?" Students volunteered that the first two columns were increasing and the last was decreasing. A good initial look at the data completed, the instructor moved on to asking, "How does the data change?" The instructor pushed them to examine relative densities as a way to fairly compare the cultures Next, the instructor suggested it might be easier to compare the cultures if the initial values were the same. The corresponding transformation produced the table shown in Figure 2. Then students calculated the relative change and a pattern emerged to lead to an exponential function.

Methods

The study was carried out using survey methods using an existing instrument (Powers, Bright, & Bugaj, 2010). There were two types of surveys: a large online pre- and post-survey and four smaller in-class surveys. The questionnaires focused on many aspects of the adaptation of instructional technology in mathematics courses, but specifically the questionnaire asked three questions related to course cohesion:

1. I expect pre-class multi-media materials to prepare me to participate in class activities (group discussion, problem solving, etc.)
2. As a result of the out-of-class material, I expect to be confident in my understanding of the concepts that each module covered.
3. I expect the in-class activities to be clearly coordinated with the pre-class material.

These questions were given on a 5-point Likert scale and the verb tense was changed from the pre- to the post-questionnaire. Results of the survey were analyzed using Qualtrics descriptive statistics. Students were also given the opportunity to provide open-ended feedback on the surveys and qualitative responses were coded accordingly.

The in-class surveys asked the following questions related to the prelectures:
1. Did you complete the prelectures this week? (Circle one) All Some None
2. Were the lectures related to the prelectures this week?
3. If they were related, give an example of something from a prelecture that you felt was useful in a lecture.

Results

The pre-/post-surveys online were completed by 58 and 20 students, respectively. The low response rate on the post-survey might be due to its administration being close to finals. Table 1 shows student responses to the pre-/post-surveys respectively.

Table 1. Results from the online surveys.

Question (5-point Likert scale)	Pre- Mean (SD)	Post- Mean (SD)
I expect pre-class multi-media materials to prepare me to participate in class activities (group discussion, problem solving, etc.)	3.67 (0.95)	3.60 (0.67)
As a result of the out-of-class material, I expect to be confident in my understanding of the concepts that each module covered.	3.16 (1.03)	3.10 (1.94)
I expect the in-class activities to be clearly coordinated with the pre-class material.	3.88 (0.89)	3.90 (0.97)

Of the 20 students who took the post survey, 65% reported that course design and instruction met or exceeded their expectations. Table 2 shows the results from the questions related to the online modules. There was no statistically significant difference in students' perceptions of the alignment between in- and out-of-class materials. This suggests that they found the materials helpful in maintaining instructional coherence. Although the mean response on the second question did not change significantly from pre- to post-survey, the standard deviation nearly doubled. The reasons for this were explained by the open-response questions about aspects of the course. Students indicated that they either did not need the additional help and so found little use in the prelectures or they really appreciated the introduction to lecture material.

Table 2. Results from the in-class surveys

	Survey 1 (n=59)	Survey 2 (n=62)	Survey 3 (n=53)	Survey 4 (n=55)
Completed all the prelectures	85%	82%	72%	81%
Prelecture was related to the lecture	85%	74%	85%	96%

Our main purpose was to investigate the efficacy of a method of applying the flipped classroom model in a meaningful and coherent way that is based in theory of mathematics teaching and learning. Analysis suggests that a focus on cognitive obstacles relative to the in-class material improves coherence of instruction in the flipped classroom model. In this way, flipped classroom models can provide a way of addressing cognitive obstacles in addition to being an alternative and cost-effective option. We see that "technology is here to transform thinking, and not to serve as some prosthetic device to prop up old styles of pedagogy or curriculum standards" (Hegadus & Moreno-Armella, 2009, p. 397). Future research needs to check this out with larger-n study. Qualitative responses to the in-class surveys also indicated students were making connections between current material and past mathematical material leading toward curricular coherence. More research is necessary to assess curricular coherence.

Questions for the Panel:

How else, besides survey questions, can we measure students' perception of coherence of a course? Are there alternative ways to operationalize coherence, besides relying on student perceptions?

What other cognitive obstacles do students bring to differential equations?

How else have instructors/researchers implemented flipped classrooms successfully?

References

Andrews, T. M., Leonard, M. J., Colgrove, C. A., & Kalinowski, S. T. (2011). Active Learning *Not* Associated with Student Learning in a Random Sample of College Biology Courses. *CBE -- Life Sciences Edition, 10*, 394-405.

Bingolbali, E., & Ozmantar, M. (2009). Factors shaping mathematics lecturers' service teaching in different departments. *International Journal of Mathematical Education in Science and Technology, 40*(5), 597-617.

Bingolbali, E., Monaghan, J., & Roper, T. (2007). Engineering Students' Conceptions of the Derivative and Some Implications for Their Mathematical Education. *International Journal of Mathematical Education in Science and Technology, 38*(6), 763-777.

Bowers, J., & Zazkis, D. (2012). Do students flip over the flipped classroom model for learning college calculus? *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

Czocher, J. A., Tague, J., & Baker, G. (2013). Where does the calculus go? An investigation of how calculus ideas are used in later coursework. International Journal of Mathematical Education in Science and Technology, 44 (5), 673-684.

Herscovics, N. (1989). Cognitive obstacles encountered in the learning of algebra. In S. Wagner & C. Kieran (Eds.). Research issues in the learning and teaching of algebra. (pp. 60-86). Reston, VA: National Council of Teachers of Mathematics.

Monk, G. (1992). Students' understanding of a function given by a physical model. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy*. *MAA Notes, Volume 25*. Washington DC: Mathematical Association of America.

Rasmussen, C. (2001). New directions in differential equations: A framework for interpreting students' understandings and difficulties. *The Journal of Mathematical Behavior, 20*(1), 55-87.

Thompson, P. W. (1994). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*, 151-169.

Zandieh, M. (1997). *The evolution of student understanding of the concept of derivative.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9805485)

	Temperature		
Time in hrs.	25°C	35°C	45°C
0.0	6.68	2.67	7.88
1.0	7.62	3.87	7.33
2.0	8.68	5.61	6.86
3.0	9.89	8.09	6.38
4.0	11.22	11.72	5.99
5.0	12.83	16.98	5.52
6.0	14.56	24.59	5.20

Figure 1. Table from Baker (2012) (p. 2)

	Temperature		
Time in hrs.	25°C	35°C	45°C
0.0	1.00	1.00	1.00
1.0	1.14	1.45	0.93
2.0	1.30	2.10	0.87
3.0	1.48	3.03	0.81
4.0	1.68	4.39	0.76
5.0	1.92	6.36	0.70
6.0	2.18	9.21	0.66

Figure 2. Table from Baker (2012) (p. 2)

DISAMBIGUATING RESEARCH ON *LOGIC* AS IT PERTAINS TO ADVANCED MATHEMATICAL PRACTICE

Paul Christian Dawkins
Northern Illinois University

Many consider logic a hallmark of mathematical practice and an integral part of proof-oriented mathematical instruction. This is true of the term logic *whether it refers to a domain of mathematical study or to aspects of reasoning, but I claim that these formalized and psychological senses of the term must be carefully distinguished in mathematics education research. In the course of identifying how the abstraction criterion has been misapplied across various types of logic in psychological and mathematics education research, I outline a framework for the disambiguation of the range of research constructs referred to as* logic. *By distinguishing the types of logic pertinent to mathematics education instruction, I hope to provide a language by which future research can better specify the constructs they investigate. Clearer research constructs should help the community to understand the role various logics play in students' apprenticeship into the practices of advanced mathematics.*

Key words: Logic; Reasoning; Abstraction; Conditionals; Proof

People often justify mathematics instruction because it develops logical reasoning skills (Gonzalez & Herbst, 2004; Inglis & Simpson, 2008, 2009). Unfortunately, many studies demonstrate that students' untrained reasoning differs from the conventional logic of mathematics (e.g.- Dubinsky & Yiparaki, 2000; Epp, 2003; Hoyles & Kuchemann, 2002), suggesting that logic poses a barrier to success in advanced mathematics. The exact role of logic in the transition to proof-oriented mathematics is debated because mathematicians explicitly use formal logic very little (Hanna & De Villiers, 2008; Thurston, 1994). As many universities across the United States introduce "bridge courses" meant to help students succeed in proof-oriented mathematics classes, a debate recurs regarding whether and how formal logic should be taught. Furthermore, faculty often assess students' proofs in terms of logical validity, but little is understood about how and whether reasoning about logic itself played a role in producing the proof. Like many other aspects of the complex activity of proving, certain aspects of logic seem to be necessary, but by no means sufficient, conditions for successful enculturation to proof practices. In this paper, I contend that ambiguity among uses of the term *logic* create or at least exacerbate much of the confusion around the importance, role, and instruction of *logic* in proof-oriented mathematics. Synthesizing studies from multiple disciplines, I propose a framework for disambiguating various uses of the term *logic* to help future investigations to clarify the nature and intent of the constructs under investigation. In particular, I motivate the need for disambiguation via the criterion of *abstraction*.

Disambiguating Logic: Important Dimensions of Variation

An important question in all cognitive work in mathematics education is whether the phenomena being described are *prescriptive* or *descriptive*. Mathematics education must attend to both because student-centered paradigms of instruction require sensitivity to student reasoning as it occurs (descriptive), but instruction requires learning goals that privilege particular modes of activity (prescriptive). While the mathematics education community has made great strides toward appropriately investigating the constructed rationality of student activity, problems arise when researchers appropriate the same language or representation system to describe both constructs creating ambiguity or conflation. This pitfall particularly plagues research on *logic* because of its relation to (deductive) *reasoning*. Common parlance interchanges descriptors like "logical" with "reasonable" (meaning reasoning deemed appropriate), implying a clear standard of assessment. Furthermore, even logicians infer some connection between prescriptive logical systems and reasoning. For example, when the logician Parsons critiqued Piaget's use of propositional calculus notation, he said, "The theorems (tautologies) of truth-functional logic are not always true under [Piaget's model of psycho-logic]… What fails to make logical sense can hardly make psychological sense in a study of intellectual development" (quoted in Bond, 1996, p. 180-181). However, logicians widely acknowledge there are multiple possible "logics" none of which appears completely able to accommodate the range of common language or reasoning (see Durand-Guerrier, 2003).

It is problematic to claim students do not reason according to "logic" (by whatever standards) since it suggests they are "illogical", introducing connotative baggage. Overton (1990) even claims, "Deductive propositional and predicate logics ultimately are formalizations of the commonsense, correct deductive arguments that people engage in on a day to day basis" (p. 5). So, it would seem that to contradict the classical logics is either to be nonsensical or simply incorrect. However, some instantiations of logical reasoning render this a problematic stance. Consider the research findings related by Markovits (2004):

> "Young children do have problems making [the modus ponens] inference when contrary-to-fact premises are used… For example, if given premises such as "if it is raining, then the grass is dry," young children show a strong tendency to conclude that "it is raining" implies that "the grass is wet." However, when these children are given some external support… they consistently make the logical modus ponens inference." (p. 323-324)

While I shall elaborate later why the conclusion "the grass is dry" is considered prescriptively *logical*, it seems odd to lament when early elementary children infer that the grass is wet when it rains. The expectation that children should blindly adopt contrary-to-fact hypotheses seems *unreasonable* (in the everyday sense of that term). I argue that this research paradigm represents a misapplication of a formalized system (the prescriptive system) to everyday reasoning. Clearly the logic of the child's reasoning (descriptive) was different from that prescribed, but I would not deem it inappropriate.

The Prescriptive Logic of Mathematics

Formal mathematics, such as Euclid's *Elements,* served for many centuries as a paradigm of good reasoning. Modern formal logic similarly arose among mathematicians. Such formal logical tools are formalizations and abstractions that have been so successful in assimilating mathematical argument and language that many mathematicians view those formalizations as obvious or natural. For instance, Katz and Starbird (2013) stated, "The whole of mathematics… is merely a refinement of everyday thinking. Proving theorems [is] not a different way of thinking—it is merely a refinement of clear thinking" (p. 2). While some might question this claim based on empirical observation of non-mathematicians (e.g.- Evans, 2005; Oaksford, Chater, & Stewart, 2012; Markovits, 2004), mathematical reasoning seems to be translatable into the formalisms of propositional and predicate logics. Mathematics educators are concerned with mathematicians' reasoning patterns because they should constitute a prescriptive system for mathematics instruction, especially in advanced mathematics classes where students should be acculturated into the mathematics community. If mathematicians reason in ways compatible with formal logic, then that logic will provide a prescriptive system for mathematics instruction.

Unfortunately, the relation between mathematicians' reasoning and formalized logic is less clear. The most common representative of *logical* reasoning in research is the conditional ("if…then…") and many assume that mathematicians reason with the *material conditional* (MC) in which "if p, then q" is equivalent to "not p or q." Because logicians assume MC logic is *abstracted* from particular content, there need be no relevance (Piaget & Garcia, 1991) relation between p and q. Inglis and Simpson (2004) administered a standard psychological test for conditional reasoning to mathematicians and found that while they reasoned differently from non-mathematicians, their performance did not coincide with the MC. However, this may reflect more on the Wason card task (see Evans, 2005, for a summary) as a research tool rather than the logical reasoning it is assumed to measure. More strikingly, Weber and Alcock (2005) provided mathematicians with conditional statements that were true by MC standards, but with no clear mathematical link between the antecedent (p) and the consequent (q). Mathematicians did not affirm such claims. The authors proposed that mathematicians instead assessed the truth of conditional statements using a *warranted conditional* (WC) in which the claim was true only if they knew warrants by which the consequent could be proven from the antecedent. So mathematician's interpretation of conditionals relies less on truth-function (meaning it is derived solely from the truth-values of the component propositions) than on *provability*, which intimately relates to the semantic content of p and q.

So, the (descriptive) logic of mathematics students' reasoning differs from mathematicians' (descriptive) logic (Epp, 2003), but empirical studies also show mathematicians' reasoning is not fully compatible with (prescriptive) formalized logic. Despite this, in transition-to-proof courses mathematicians often introduce MC-based truth-table analysis as a prescriptive model for mathematical reasoning (Selden, 2012).

Key Assumptions Derived from Formal Systems

Based on the previous studies, I find it difficult to maintain that formal logic is an appropriate prescriptive system for the logic of mathematical or non-mathematical reasoning, but this continues to be an attractive hypothesis to many. A number of psychologists have spent extensive time attempting to confirm the existence of some innate logical (in the formal sense) mechanisms in human reasoning (Braine, 1990; Oaksford et al., 2012). While I agree with Stylianides and Stylianides (2008) that mathematics educators should attend to psychological research about deductive reasoning, I urge caution about the assumptions behind such research. The way many psychologists test "deductive reasoning" is via the Wason card task, but the afore-mentioned studies suggest this test does not capture the logic of mathematicians' reasoning. One problem is that psychologists' definition of "deductive reasoning" embeds the formal logical criterion of *abstraction* into their research assumptions (Evans, 1982; Evans & Feeny, 2004). *Abstraction* asserts that the logic of an argument is independent of the semantic meanings of the statements being reasoned about (see Durand-Guerrier et al., 2012; Overton, 1990). For instance, in the research Markovits (2004) described, children's reasoning is deemed *logical* if it ignores their knowledge of rain (semantic meaning) and relies solely on the given assumption ("If it is raining, then the grass is dry"). More recently, psychologists questioning these assumptions have proposed alternative framings for the logic of naïve and developmental reasoning such as probabilistic reasoning (Evans & Feeny, 2004), systems of pragmatics (Markovits, 2004), or the logic of meanings (Piaget & Garcia, 1991).

As early as Aristotle and more prominently in the 20th century, logicians used *abstraction* to distinguish logic from rhetoric and to identify undue assumptions. Logicians embody abstraction through a notation of linguistic variables such as p, q, or P(x). Hilbert famously emphasized the logical validity of his geometry axioms saying, "Instead of treating 'points', 'straight lines' and 'planes', one must always be able to discuss 'tables', 'chairs' and 'beer-mugs'." (quoted in Mariotti & Fischbein, 1997). I think it is important that this provides a heuristic for assessment rather than a means of axiomatizing, but the criterion of abstraction seems clear. Hilbert's stance should not be overstated, since mathematicians in Weber and Alcock's (2005) study did care about mathematical connections between the contents of a conditional claim. Semantic content is clearly important, seeing as it is easier to affirm modus ponens inferences when the propositions are abstract (p and q) than when they are particular contrary-to-fact claims (as in the rain example above). Since formalized logic distinguishes *validity* (abstract form of argument) and *truth* (correspondence with understanding or experience), many researchers have assumed truth (and thus prior knowledge) should be irrelevant in logical reasoning.

Problems of Communication and Interpretation

The notations that embody the abstraction criterion cause confusion about the meaning of *logic* in descriptive research (Overton, 1990; Bond, 1996). Piaget and colleagues (Inhelder & Piaget, 1958; Piaget & Garcia, 1991) sought to describe "the logic of meanings" children use to solve tasks. In line with Piaget's constructivism, this *descriptive* project tied meaning to students' action rather than abstract concepts or propositions. He used formal logical notation to classify and distinguish the reasoning of action, much as he used algebraic group language elsewhere. Piaget (1950) explained this practice when he cautioned those who wanted to make "thought the mirror of logic" saying, "reverse the terms and make logic the mirror of thought, which would restore to the latter its constructive independence" (p. 30). Unfortunately, even prominent psychologists completely misunderstood his use of such notations (Overton, 1990) assuming he claimed that "adult human reasoning was inherently logical", by which they mean consistent with the MC (Evans & Feeny, 2004). They assumed the abstract notation implied abstractness in the reasoning it modeled.

Similar problems can appear in mathematics education research. For instance, Stylianides, Stylianides, and Phillipou (2004) investigated students' reasoning about contrapositive arguments. One hypothesis they presented in an argument by contraposition stated, "If Costas suffered from pneumonia, he would have high fever" (p. 139). The authors rejected the response of one student who contested, "Perhaps fever is not the only symptom. Therefore, the conclusion is wrong" (p. 147). The authors' justification for their critique assumes the given hypothesis is true, consistent with mathematical convention, in which case the students' counterargument is invalid. I contend that the students' argument is appropriate according to everyday pragmatic conventions: hypotheses may be questioned and there may exist cases of pneumonia with no fever. The two interpretations directly correspond to *truth* (the student's concern) versus *validity* (the authors' concern). The difference concerns the pragmatics by which people determine how statements should be interpreted and the standards of discourse that are applicable in a given situation (part of what Stenning and van Lambalgen, 2004, call *reasoning to an interpretation*).

Pragmatics aside, Stylianides et al. (2004) made a stronger claim with regard to the aforementioned student argument: "These students did not manage to decontextualize the statements and assess the necessity of the conclusion in purely logical terms" (p. 147). In this way, the authors not only compared the students' claim against the outcome of their own formalized logical reasoning, but also claim that logical reasoning should involve some level of abstraction away from semantic meaning. Thus the *abstraction* used by logicians to formalize rules of logic becomes imbued with a prescriptive cognitive aspect. I question this conflation of constructs and insist researchers must distinguish whether *logic* refers to the actual process of reasoning *psychologically*, or to the *formalized* system used to develop rules for validity (i.e.- heuristics to assess the outputs of reasoning). Stylianides et al. (ibid.) assumed the *abstraction* of the latter applies to the former by insisting logical reasoning must in some sense ignore semantic content. Many psychological researchers make similar assumptions (Oaksford et al., 2012).

I claim instead that *abstraction* must be considered a useful *convention* of *formalized* logic rather than a prescriptive *assumption* about logical *reasoning*. To justify this, imagine a mathematician expositing a basic axiomatic proof to a class of students. Suppose the mathematician "applies" a conditional axiom (e.g.- "If m and n are numbers and their successors are such that $m'=n'$, then $m=n$") in what could be called a *modus ponens* inference. Would it be appropriate to say the professor used *modus ponens* to draw an inference (as is common in psychological research)? If logic (including the *abstraction* condition) refers to reasoning, then it is unlikely at best. This would mean that she somehow ignored her knowledge of the natural numbers and divorced the formal notion of "successor" from the counting it models (a rather fantastic claim) and depended upon the conditional axiom to deduce a conclusion. As Detlefsen (2008) argues, mathematical proof is more often considered *translatable* into formal logic rather than existing *a priori* in that representation system. I posit the mathematician has trained herself to *conform* her semantic reasoning (in this case about numbers) to the syntactic standards of logic (what I called heuristics for assessment), or more specifically to the *conventions* of mathematical proof.

I find it much more reasonable to characterize abstraction as a useful *convention* adopted by mathematicians for attaining valid argument. I also hold that the Platonism of many mathematicians (Rotman, 2006) implies that they value validity precisely because it produces true results (what logicians call *soundness*). Psychologically, the axiom began as a *formal articulation* of the mathematicians' knowledge of natural numbers based on years of *pre-formalized* experience. The essential mathematical practice of abstraction involves shifting focus from the particular objects of mathematical activity to their relationships and properties, which are the bedrock of proof. Unfortunately, mathematicians may not recall the reorganization process engendered by proof practice, to the extent of conflating axiomatic structure (formalization) with the epistemic foundations of numbers (psychological). Similarly in logic, some claim students "misinterpret the meaning" of *if* (Anderson, 2010, p. 289) suggesting that the MC is the right interpretation of the linguistic form rather than the product of a particular formalized game of reflecting on arguments (what I call "reasoning about logic").

A Basic Framework for Disambiguating Types of Logic

In summary, *logic* can be prescriptive or descriptive and it can concern psychological processes or formalized systems. Combining these two dichotomies, there are at least four distinct categories of meaning for the term *logic* as presented in Table 1. Prescriptive-formalized (Type 1) logic entails the standard propositional and predicate logics developed by Frege, Russell, Witgenstein, and Tarski (see Durand-Guerrier, 2008). As previous research shows (e.g.- Epp, 2003; Evans & Feeny, 2004), the logic of mathematics students' reasoning (Type 3) differs from Type 1 logic in multiple ways. As I claimed above, mathematicians' reasoning (which constitutes Type 2 logic for proof-oriented mathematics education) is not identical to Type 1 logic either (see also Nickerson, 2004), but rather mathematicians learn to *conform* their proofs to these standards up to some level of contextual tolerance (Reid, 2011). Regarding Type 4 logic, mathematicians and philosophers in proof theory and related fields continue to extend formalized systems in new prescriptive and descriptive directions (Piaget & Garcia, 1991, reference the work of Anderson & Belnap, 1975). However, such work is not fully necessary for mathematics education since Type 1 logic articulates the standards to which Type 2 reasoning conforms.

	Formalized system	Psychological process
Prescriptive	(Type 1) Propositional/predicate	(Type 2) Conformal
Descriptive	(Type 4) Non-standard	(Type 3) Naïve/pragmatic

Table 1: Categories of meaning for *logic* in research

Conclusions and Future Directions

To restate my previous argument in this new terminology, distinguishing these categories of logic is important because *abstraction* is a useful convention for Type 1 logic, is a goal for Type 2 logic, and can only be problematically applied to Type 3 logic. I hypothesize that greater attention to clearly identifying the constructs under investigation will improve research on this important component of proof practice. Specifically, future research may investigate how mathematicians *conform* their Type 2 reasoning in action. While the work of many mathematics educators (Durand-Guerrier, 2003, 2008; Epp, 2003; Roh & Lee, 2011) provide rich activities for making students more consciously aware of the conventions of Types 1 and 2 logic, the research community must continue to clarify research methods on Type 3 logic to properly implement and understand the learning engendered by these activities. Being that Type 1 logic arose as a formalization of Type 2 activity, I also conjecture that an instructional sequence for reinvention of logic and validity could yield fruitful insights about how to help students transition Type 3 reasoning toward Type 2 reasoning, which I claim is a primary goal of proof-oriented instruction.

References

Anderson, J.R. (1990). *Cognitive Psychology and Its Implications (3rd ed.)*, New York City, NY: Freeman.

Bond, T. (1996). Piaget and Measurement. In Smith, L. (Ed.) *Critical Readings on Piaget* (pp. 178-208). New York, NY: Routledge.

Braine, M. D. (1990). The “natural logic” approach to reasoning. In Overton, W. (Ed.) *Reasoning, necessity, and logic: Developmental perspectives* (pp. 133-157), Hillsdale, NJ: Lawrence Erlbaum Associates.

Detlefsen, M. (2008). Proof: Its nature and significance. In Gold, B. & Simons, R. (Eds.), *Proof & other dilemmas: mathematics and philosophy* (pp. 3-32), Washington, DC: Mathematics Association of America.

Dubinsky, E. & Yiparaki, O. (2000). On student understanding of AE and EA quantification. *Research in Collegiate Mathematics Education, IV,* 239-289.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in mathematics*, *53*(1), 5-34.

Durand-Guerrier, V. (2008). Truth versus validity in mathematical proof. *ZDM*, *40*(3), 373-384.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S.S., & Tanguay, D. (2012). Examining the Role of Logic in Teaching Proof. In Hanna, G. & De Villiers, M. (Eds.) *Proof and Proving in Mathematics Education* (pp. 369-389). Netherlands: Springer.

Epp, S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly, 110,* 886-899.

Evans, J. (1982). *The psychology of deductive reasoning*. London: Routledge & Kegan Paul.

Evans, J. (2005). Deductive reasoning. In Holyoak, K.J. & Morrison, R.G. (Eds.) *Cambridge Handbook of Thinking and Reasoning*. (pp. 169-184) Cambridge, NY: Cambridge University Press.

Evans, J. & Feeny, A. (2004). “The Role of Prior Belief in Reasoning” In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 78-102). Cambridge, U.K.; New York: Cambridge University Press.

González, G., & Herbst, P. G. (2006). Competing arguments for the geometry course: Why were American high school students supposed to study geometry in the twentieth century?. *International Journal for the History of Mathematics Education*, *1*(1).

Hanna, G., & de Villiers, M. (2008). ICMI Study 19: Proof and proving in mathematics education (Discussion document). *ZDM-The International Journal of Mathematics Education, 40*, 329-336.

Hoyles, C. & Kuchemann, D. (2002). Students’ understanding of logical implication. *Educational Studies in Mathematics, 51*(3), 193-223.

Inglis, M., & Simpson, A. (2004). Mathematicians and the selection task. In *Proceedings of the 28th International Conference on the Psychology of Mathematics Education* (Vol. 3, pp. 89-96).

Inglis, M., & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics*, *67*(3), 187-204.

Inglis, M., & Simpson, A. (2009). Conditional inference and advanced mathematical study: Further evidence. *Educational Studies in Mathematics*, *72*(2), 185-198.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York, NY: Basic.

Katz, B. & Starbird, M. (2013). *Distilling Ideas: An Introduction to Mathematical Thinking*. Washington, D.C.: Mathematical Association of America.

Mariotti, M.A. & Fischbein, E. (1997). Defining in Classroom Activities. *Educational Studies in Mathematics, 34*(3), 219-248.

Markovits, H. (2004). “The Development of Deductive Reasoning” In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 313-338). Cambridge, U.K.; New York: Cambridge University Press.

Nickerson, R. (2004). “Teaching Reasoning” In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 410-442). Cambridge, U.K.; New York: Cambridge University Press.

Oaksford, M., Chater, N. & Stewart, N. (2012). “Reasoning and decision making”. In K. Frankish & W. Ramsey (Eds.) *The Cambridge Handbook of Cognitive Science*. (pp. 131-150). Cambridge University Press: New York.

Overton, W. (1990). Constraints on Logical Reasoning Development. In Overton, W. (Ed.) *Reasoning, Necessity, and Logic: Developmental Perspectives* (pp. 1-32). Hillsdale, NJ: Lawrence Erlbaum.

Piaget, J. (1950). *The psychology of intelligence.* New York, NY: Routledge Classics.

Piaget, J. & Garcia, R. (2013). *Toward a logic of meanings*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reid, D. (2011). Understanding Proof and Transforming Teaching. In. Weist, L.R. & Lamberg, T. (Eds.) *Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 15-30). Reno, NV: University of Nevada, Reno.

Roh, K., & Lee, Y. (2011). The Mayan activity: A way of teaching multiple quantifications in logical contexts. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 21*(8), 1-14.

Rotman, B. (2006). Towards a semiotics of mathematics. In Hersh, R. (Ed.) *18 Unconventional Essays on the Nature of Mathematics* (pp. 97-127). New York City, NY: Springer.

Selden, A. (2012). Transitions and proof and proving at tertiary level. In Hanna, G. & de Villiers, M. (Eds.) *Proof and Proving in Mathematics Education: The 19th ICMI Study* (pp. 391-420). Netherlands: Springer.

Stenning, K. & van Lambalgen M. (2004). "A little logic goes a long way: basing experiment on semantic theory in the cognitive science of conditional reasoning". *Cognitive Science, 28*, 481-529.

Stylianides, A.J., Stylianides, G.J., & Philippou, G.N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics, 55*, 133-162.

Stylianides, G.J. & Stylianides, A.J. (2008). Proof in School Mathematics: Insights from Psychological Research into Students' Ability for Deductive Reasoning. *Mathematical Thinking and Learning, 10*, 103-133.

Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*, 161-177.

Weber, K., & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics*, *25*(1), 34-51.

53

HOW DOES UNDERGRADUATES' UNDERSTANDING OF THE FUNCTION CONCEPT EVOLVE DURING THE COURSE OF A SEMESTER?

Eyob Demeke, University of New Hampshire
Vincent Mateescu, Rutgers University
Anek Janjaroon, University of New Hampshire

Functions are a crucial topic in the study of mathematics. Research has found that a lack of deep understanding of functions is one of the main reasons why students struggle in calculus (Eisenberg, 1991; Ferrini-Mundy & Graham, 1991; Lauten, Graham, Ferrini-Mundy, 1994; McDonald, Mathews, & Strobel, 2000; Monk, 1994). In light of these studies, we investigate – using traditional paper-and-pencil assessments, concept maps, and an interview – what pre-calculus students' understanding of functions is, to what extent students have a repertoire of functions at their disposal, how students' understanding evolves over a semester, and what non-traditional assessments can tell us about this understanding. We found that (1) As Williams (1998) suggested, concept map assessments do reveal something that traditional assessments do not; (2) participants have trouble giving non-examples of functions, and (3) there does not seem to be a major change in participants' understanding of functions over time.

Key words: Functions, Concept Maps, Assessment, Facets And Layers.

In most tertiary institutions in the United States, the vast majority of students who are either in the mathematics or sciences program take some form of calculus as an entry-point to undergraduate mathematics study. In the study of calculus, a rich understanding of functions is needed. In fact, the concept of functions is a prerequisite to understanding many central concepts in calculus such as the limit, derivative, and integral (Oehrtman, Carlson, & Thompson, 2008). Indeed, in all of mathematics, the topic of functions is considered to be the most important concept (O'Callaghan, 1998; Ferrini-Mundy & Graham, 1989). In light of this, some universities designed a prerequisite course to calculus that aims at strengthening students' understanding of the concept of functions.

There are a plethora of studies on students' understanding of functions. For an extensive treatment of the literature on this subject, we refer to Leinhardt et al. (1990). The literature shows that students frequently struggle to understand functions, as evidenced by a number of their fundamental misunderstandings of the concept. In order to be successful, students must understand functions as general processes that accept an input and produce an output, and they must be able to think about the change of the output as the input changes and its rate of change (Oehrtman, Carlson, & Thompson, 2008). The goal of any precalculus class is to help students overcome the types of misconceptions that they arrive with and develop the strong notions of function and variation that they will need in order to be successful, yet little is known about how students' understandings change over the course of such a class. Multiple studies have noted the inadequacy of traditional assessments as tools for describing student understanding (Carlson, Oehrtman, & Engelke, 2010; Williams, 1998). Thus, we investigate the following research questions:

- How does the understanding of functions of undergraduate students evolve over the course of a semester?
- How do different forms of assessment give access to and constrain our ability to investigate students' understanding?

Theory

The perspective of learning assumed in this paper is Piagetian constructivism, the idea that knowledge is constructed, rather than acquired. This theory of learning is based on the fundamental notion that knowledge is not passively received; rather, it is an adaptive activity that is actively built up by the cognizing subject (von Glasersfeld, 1995). Piaget argued that learning occurs as a consequence of the assimilation of new knowledge with existing knowledge. Assimilation is the process by which the individual takes in new data. The process of assimilation is followed by accommodation, which is the process by which the individual makes some modification in his/her cognitive structure (Piaget, 1970).

DeMarois and Tall (1996) described a conceptual framework for students' understanding of functions that characterizes understanding of functions in terms of *facets* and *layers*.

- *Facets* are approaches (breadth) one can take to understand the concept of function. DeMarois and Tall (1996) characterize *facets* in a *horizontal* sense.
- *Layers* describe the extent (depth) to which one examines the concept in question in any one *facet*. DeMarois and Tall (1996) characterize *layers* in a vertical sense.

Their take on the function concept employs eight different facets – *verbal* (spoken), *written*, *kinesthetic* (enactive), *colloquial* (informal or idiomatic), *notational conventions*, *numeric*, *symbolic*, and *geometric* (visual) aspects. DeMarois and Tall (1996) acknowledged that these facets are not intended to be neither independent nor exhaustive ways of approaching the concept of function. Their version of the function concept also includes a superficial layer (*pre-action*), *action*, *process*, *object*, and core layer (*procept*) as the five different levels.

Previous research on student understanding of functions

In general, instruction on functions has focused on procedures without deep understanding, and it has been insufficient. In the section that follows, we describe some of the most important findings about student difficulty and relate them, as appropriate, to the DeMarois and Tall (1996) framework. In general, students have (at best) an action-view of function (DeMarois & Tall, 1996; Oehrtman, Carlson, & Thompson, 2008). As a result, they have difficulty distinguishing between algebraically defined functions and equations (Carlson, 1998) and believe that all functions should be able to be defined by one algebraic formula (Breidenbach et al., 1992), which, in turn, leads them to believe that all functions are "nice." This type of conception results in believing that a piecewise function represents multiple, distinct, functions and that there is a single unique formula that defines any particular function (Oehrtman, Carlson, & Thompson, 2008).

When students develop a more abstract understanding, such as that at the *object* or *procept* layer, they are able to consider functions as a process that may be reversed (resulting in an inverse function), consider composition of functions as an operation on two objects (as opposed to mere algebraic manipulation) and consider, simultaneously, the relationship between all possible input and output values. Such a view results in a nuanced understanding where students are able to consider properties of the functions and think about structure (Williams, 1998). In general, students are less likely to think about function families than experts (Williams, 1998), and, even if they do, are likely to assume that any function whose graph looks like a line is linear and any function whose graph is u-shaped is a quadratic (Schwarz & Hershkowitz, 1999). That is, they categorize into function-families based upon surface characteristics, as opposed to the structural thinking in which experts engage.

Research methodology

This section will describe the ways through which our group collected, and then analyzed, our data. Our choices of both population and sample were appropriate. The population consisted of students from a precalculus class at a research university. From our population, our group eventually selected three participants to participate in a detailed case study. This number is appropriate because it allowed for generalizations to be made about the larger population, but also, because they exhibited diversity of opinion about an item on a background. In particular, their opinions varied when asked to evaluate the plausibility of the statement "Mathematics is mostly computational."

We collected data across three "phases" during the semester of the study. In phase I, our group administered a preliminary questionnaire and the traditional assessment. In phase II, our group had the participants re-take the same traditional assessment, explain what a concept map is, give an example of a concept map (that had "numbers" as the central topic rather than "functions"; we were careful to choose a topic different from the one on which students would make their concept maps), and create their own concept maps for "function". Finally, in phase III, our group administered the same traditional assessment a third time, the concept map assessment a second time, and an interview – the purpose of the interview was twofold; first, to clarify remarks on either the traditional or concept map assessment that our group found confusing, and second, to test each participant for a particular kind of understanding that neither the traditional nor concept map assessment can detect: kinesthetic understanding (DeMarois & Tall, 1996, p. 4). Taken together, these three phases occurred in September, November, and December, respectively. The data we collected during each phase allowed us to capture snapshots of the students' understanding of functions at various points in the semester and, read together, to describe the changes in their understanding over the course of the semester.

Our hypotheses were as follows. First, we expected the scores on the traditional assessments to increase over time for all participants. Second, we hypothesized that the participants' ratio of number of valid links to number of total links will increase with each passing phase. In particular, we expected the number of both valid links and total links on the concept maps to increase for all three participants. Finally, we expected that students expand their repertoire of examples of functions. We acknowledge these hypotheses as part of our epistemological beliefs about qualitative research and note that they influenced our approach to the task of data analysis. We were implicitly testing these beliefs against the data, and acknowledge them to guard against confirmation bias (Creswell, 2007).

In terms of our data analysis, in the first phase we drew on the diagram in the DeMarois and Tall (1996, p. 4). In particular, we applied it to the participants' traditional assessments by being on the lookout for key words on the latter that hinted at certain layers. For example, the last item on the traditional assessment gave students two graphs, asked them whether or not they were functions, and required students to explain their reasoning. Based on their explanations, we assigned a layer to certain facet – *geometric* – that we held constant. If the participant appealed to the vertical-line test, then we interpreted that as evidence that the participant was aware of a *process* (procedure) of testing whether or not a graph depicted a function; hence, we assigned the layer of *process*. We assigned the *procept* layer if the student appealed to the definition of function – specifically, if the student mentioned that every input yields a *unique* output. The word *unique* was important, for our group would have assigned the *object* layer had the student left it out.

We also devised a framework to analyze our participants' concept maps. Each one of us kept a tally of the number of mathematically valid and invalid links or connections for each participant's concept map. In the event where there was disagreement amongst ourselves whether or not a link is mathematically valid, we asked the participants during the interview for further clarification on their concept map. Otherwise, we then took the ratio of the number of valid connections to total connections in each concept map. In addition, each of us wrote a summary of any understanding that could be inferred from participants' concept maps. The summary included information related to (1) definition, (2) examples and types, (3) properties, and (4) operations of functions.

Results and discussion

This section will discuss three findings about the understanding of functions of one particular student, Epsilon (we gave pseudonyms to each participant). We chose to present data from Epsilon rather than all three students because we think that his results are representative of the remaining two. The first finding was in the evolution of Epsilon's understanding of the term function. In Phase I, Epsilon's response was that a function "is a rule." This answer was incomplete because Epsilon did not go into detail about what the rule involved, perhaps suggesting that he has memorized this idea. As a result, we classified his answer as a *pre-action*. In Phase II, Epsilon provided a more complicated answer: a function "is a rule… when you plug in a independent [sic] value you get another individual value out". Here, Epsilon hinted at the process of plugging in a number and obtaining a corresponding second number. Thus, we labeled this answer as the *process* layer. In Phase III, Epsilon's answer became even more sophisticated. He described a function as a "rule in which for every input there is a unique and individual output". From this answer, we believe that Epsilon identified a function as an *object* in its own right. However, not all features of his traditional assessment underwent improvement over time. For example, he kept providing examples of linear functions when asked to provide examples of functions and non-functions. Though this and other misunderstandings existed across the phases, not only did his conception of what a function is improve (as demonstrated earlier in this paragraph), but also, his percent score for Phase III was higher than that for the two preceding phases. Thus, Epsilon did somewhat better in Phase III of the traditional assessment than he did anytime earlier in the semester.

The second finding was that Epsilon's concept map evolved somewhat from Phase II to Phase III in regards to the mathematical understanding displayed therein (please see the appendix). This was true despite the ratio of valid connections to total connections decreasing from one phase to the next (as was the case for all three participants) – in Phase II, Epsilon had 18 valid connections out of 22 total, and in Phase III, he had 17 valid connections out of 26 total. This decrease was noteworthy and merits further comments, which appear later on. First, Epsilon gave notational references to different kinds of functions in Phase II, referring to rational functions as $(f/g)(x)$, inverse functions as $f^{-1}(x)$, and so forth. This changed in Phase III, with

Epsilon giving explicit examples of each kind of function. For instance, he gave $y = x^2 + x + b$ as an explicit example of a quadratic function. In fact, Epsilon gave many more examples of functions in Phase III than he did in Phase II. This was a sign that his awareness of different families of functions had become richer from one phase to the next. A second feature of Epsilon's concept maps was the presence of composition of functions as an example of an operation on functions that one can perform. This was noteworthy because Epsilon was the only student out of the three participants who recognized function composition in his concept maps. But, curiously enough, Epsilon only recognized function composition during his Phase II concept map. Thus, it is possible that Epsilon (1) felt that it was not important to include it in his Phase III concept map, (2) ran out of time, or (3) forgot to include it altogether. A final feature of Epsilon's concept maps is his labeling of certain families of functions as being positive or negative. To sum up, Epsilon's understanding of functions as shown on the concept maps became somewhat more sophisticated over time in terms of the presence of more examples and types of functions.

Our third finding was that Epsilon's interview was consistent with how he did on the traditional and concept map assessments. Epsilon said that his labels were in reference to positive and negative "slope". He then gave an example involving the function $f(x) = x^2$, which he then correctly graphed as a parabola with vertex at the origin that opened upward. Epsilon demonstrated that if he inserted a minus sign before the x^2, that would cause the parabola to reflect over the x-axis. He then immediately recognized that his choice of example was not ideal, as it was appropriate for reflections of functions (more specifically, transformations of functions) than it was for slope. This was an indication that Epsilon could not think of an appropriate function that would go better with his comment. Another feature of Epsilon's interview was his description of a function. He used an example of a specific function, $f(x) = 2x$ (that Epsilon used a linear function to illustrate his point was typical among the three participants, who would always provide this kind of function as an example of a function in general terms), and said that 1 maps to 2, 2 maps to 4, and so forth. This was an indication that Epsilon understood functions in a more tabular, numeric sense. Thus, one could appropriately label his understanding of functions under the *numeric* facet. Finally, when responding to a question about what a function was, one of the first things that Epsilon said was that a function is a "rule". Epsilon used this word in all three versions of his traditional assessment, thus establishing a connection between his traditional assessment and interview. Thus, Epsilon's responses in his interview were consistent with the answers he provided on the traditional and concept map assessments.

To sum up, in the final analysis of concept maps for each participant across all the phases, the ratio of number of valid links present to number of total links present *decreased*. However, this did not tell the full story, for at the same time, the total number of links for all three participants *increased*. This was an indication that even though more misunderstandings were present in the Phase III concept maps, we hypothesize that the students felt confident enough to take risks and see if they could expand their understanding of functions. Our data indicate two noteworthy findings about Epsilon in particular. First, his concept maps underwent some evolution regarding the richness of his understanding. Second, his responses at his interview confirmed the answers he provided on the traditional and concept map assessments. Said differently, Epsilon's understanding did not change that much over time. This finding has implications, and a discussion of some of them will follow in the next section.

Implications for further research

While it is inappropriate to draw major conclusions from a small-scale study, we do cautiously summarize our findings and interpret them in terms of their implications for the field. Yet, we see it as a logical extension of extant work, and suggest that a larger-scale follow-up is needed in order to demonstrate the robustness of these results. First, in terms of how students' understanding of functions evolves over the course of a semester of a precalculus class, we report that at the least, the students do appear to have moved from only having a very low-level layer (*pre-action*) to having an abstract way of thinking about functions (*object*). We do have concerns that the language may have been memorized rather than necessarily the student's own; this is because of related changes in the student's concept maps. In general, the students' concept maps did not show much change in understanding of functions over time; thus, it seems difficult to support the claim that students' understanding would have evolved from only a *pre-action* layer to an *object* layer without concurrent development of their understanding of other concepts related to functions. Moreover, the concept maps suggested that the students' ways of thinking about function appear to match Williams's (1998) description of students often giving surface classifications of functions as opposed to focusing on the type of structural thinking in which experts would engage. At the same time, our work shows that in spite of Williams's findings, students do classify functions into more than just linear and quadratic categories. Our students included additional families such as exponential, trigonometric, rational, and piecewise. Similarly, we showed that our students do consider piece-wise functions as entities, not as amalgamations of functions. Finally, we note that the concept map gave access to a type of thinking that none of the other assessments (traditional or interview) would have given alone. In particular, the concept map, with its completely open structure, made it possible to learn that Epsilon had developed his own quasi-structural means of categorizing functions into positive and negative, which can be understood as helpful in that it builds towards the concept of a derivative. In light of all these reasons, we suggest that researchers, in their attempt to study students' understanding of any mathematical topic, utilize assessments that target the types of understandings they anticipate, but also allow students to express their own idiosyncratic understandings that were not (and could not be) anticipated in advance.

References

Asiala, M., Cottrill, J., Dubinsky, E., & Schwingendorf, K. (1997). The development of student's graphical understanding of the derivative. *Journal of Mathematical Behavior, 16*(4), 399-431.

Asiala, M., Brown, A., Kleiman, J., & Mathews, D. (1998). The development of students' understanding of permutations and symmetries. *International Journal of Computers for Mathematical Learning*, 13-43.

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics*, *23*(3), 247-285.

Carlson, M. P. (1998). A cross-sectional investigation of the development of the function concept. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *CBMS Issues in Mathematics Education: Research in Collegiate Mathematics Education III*, *7*, 114–162.

Carlson, M., Oehrtman, M., & Engelke, N. (2010). The Precalculus Concept Assessment: A Tool for Assessing Reasoning Abilities and Understandings of Precalculus Level Students. *Cognition and Instruction, 28*(2), 113-145.

Clark, J., Cordero, F., Cottrill, J., Czarnocha, B., DeVries, D. J., St. John, D., Tolias, G., &

Vidakovic, D. (1997). Constructing a schema: The case of the chain rule. *Journal of Mathematical Behavior*, *16*(4), 345-364.

Creswell, J. W. (2007). *Qualitative Inquiry and Research Design*. Los Angeles: SAGE.

DeMarois, P. & Tall, D. (1996). Facets and layers of the function concept. In L. Puig & A. Gutiérrez (Eds.), *Proceedings of the 20th annual conference for the Psychology of Mathematics Education* (vol. 2, pp. 297-304). Valencia, Spain.

Eisenberg, T. (1991). Functions and associated learning difficulties. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 140-152). Dordrecht: Kluwer Academic Publishers.

Ferrini-Mundy, J., & Graham, K. G. (1991). An overview of the calculus curriculum reform effort: Issues for learning, teaching, and curriculum development. *American Mathematical Monthly, 98*(7), 627-635.

Lauten, A., Graham, K., & Ferrini-Mundy, J. (1994). Student understanding of basic calculus concepts: Interaction with the graphics calculator. *The Journal of Mathematical Behavior*, *13*(2), 225-237.

Leinhardt, G., Zaslavsky, O., & Stein, M. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, *60*(1), 1-64.

Monk, G. S. (1994). Students' understanding of functions in calculus courses. *Humanistic Mathematics Network Journal, 9*, 21-27.

O'Callaghan, B. R. (1998). Computer-intensive algebra and students' conceptual knowledge of functions. *Journal for Research in Mathematics Education*, *29*(1), 21-40.

Oehrtman, M., Carlson, M., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. P. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* (pp. 27-42). Washington, D.C.: Mathematical Association of America.

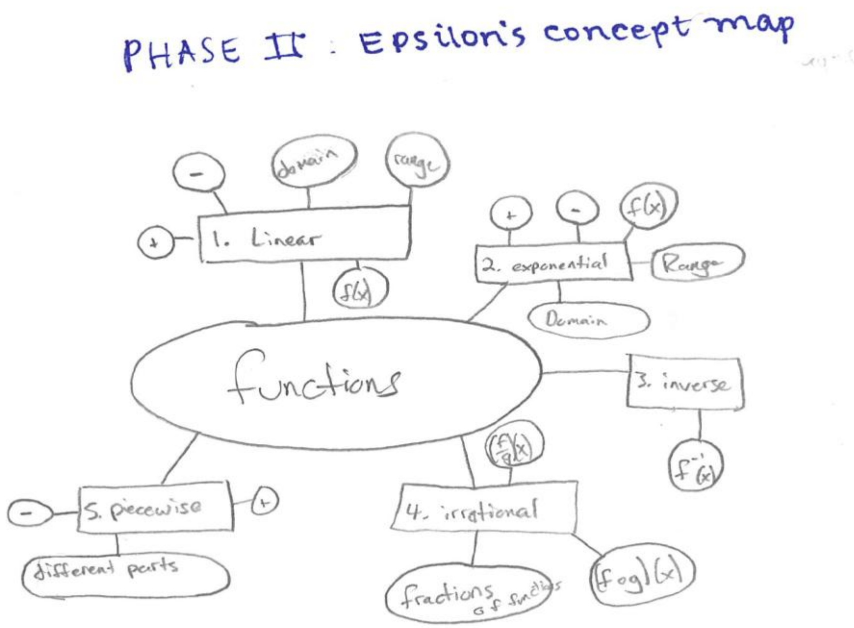
Schwarz, B. B., & Hershkowitz, R. (1999). Prototypes: Brakes or levers in learning the function concept? The role of computer tools. *Journal for Research in Mathematics Education, 30(*4), 362–389.

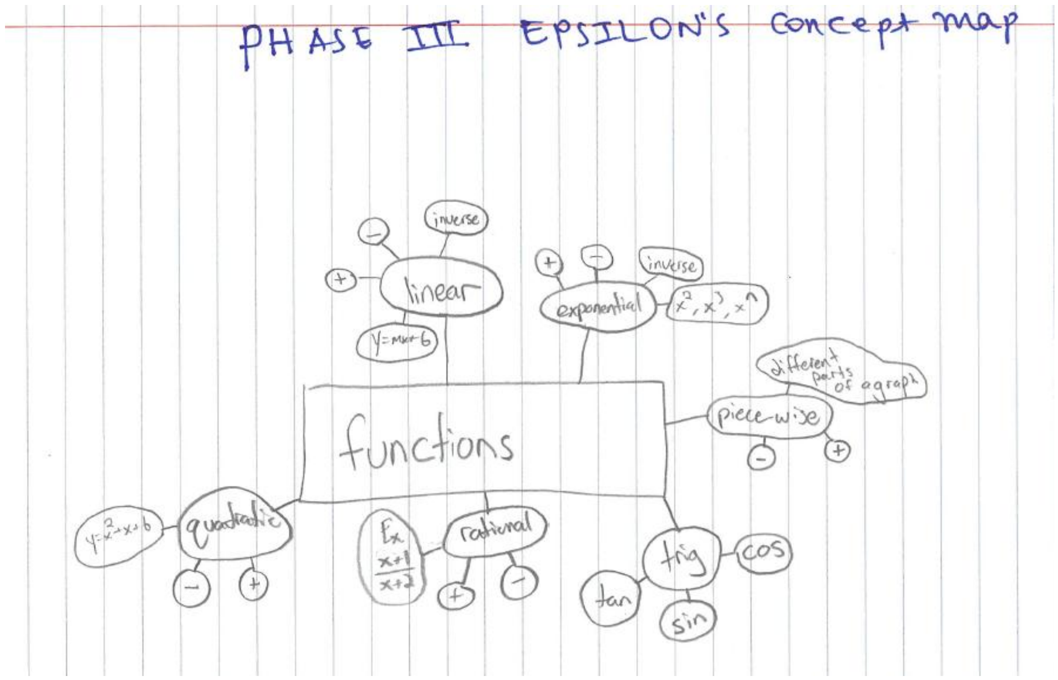
Piaget, J. (1970). *Genetic epistemology.* (E. Duckworth, Trans.) New York: Columbia University Press.

Von Glasersfeld, E. (1995). A constructivist approach to teaching. In L. P. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 3-15). Hillsdale, NJ: Lawrence Erlbaum.

Williams, C. G. (1998). Using concept maps to assess conceptual knowledge of function. *Journal for Research in Mathematics Education, 29,* 414-421.

Appendix


PHASE II : Epsilon's concept map
functions
1. Linear
domain
range
$f(x)$
2. exponential
$f(x)$
Range
Domain
3. inverse
$f^{-1}(x)$
4. irrational
$(\frac{f}{g})(x)$
fractions of functions
$(f \circ g)(x)$
5. piecewise
different parts


PHASE III EPSILON'S concept map
functions
linear
inverse
y=mx+b
exponential
inverse
piece-wise
different parts of a graph
quadratic
rational
trig
cos
tan
sin

DIFFERENTIAL PARTICIPATION IN FORMATVE ASSESSMENT AND ACHIEVEMENT IN INTRODUCTORY CALCULUS

Rebecca-Anne Dibbs University of Northern Colorado; Michael Oehrtman University of Northern Colorado

Prior formative assessment research has shown positive achievement gains when classes using formative assessment are compared to classes that do not. However, little is known about what, if any, benefits students that are not participating regularly in formative assessment gain from these assignments. The purpose of this study was to investigate the achievement of the students in two introductory calculus courses using formative assessment at the three different participation levels observed in class. Although there was no significant difference on any demographic variable other than gender and no significant difference in any achievement predictive variables between the groups of students at the different participation levels, there were significant differences in achievement on all but the first activity write-up and the final exam.

Key words: approximation framework, calculus, formative assessment

Regardless of the content area or age of participants, the effect size on most quantitative formative assessment studies is around 0.5 (Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, in press; Karpinski & D'Agostino, 2012). These studies show that classes where formative assessment is used do better on average on common summative assessments than those classes where no formative assessment is used; however, even in classes where formative assessment is used, not all students will regularly complete the formative assignments. The purpose of this study was to investigate if there were achievement differences on summative assignments in a novel calculus curriculum between students completing different numbers of formative assessments during the semester. For this paper, we will distinguish between three different low participation levels: regular, sporadic, and non-participation. Students regularly participating in the formative assessments missed no more than two formative assessments during the semester; students in the sporadic participation group completed at least one but no more than four of the 12 formative assessments in the semester, while non-participants did not complete any formative assessments.

Methods

Black and Wiliam's (2009) formative assessment framework and Vygotsky's (1987) Zone of Proximal Development (ZPD) were used as the theoretical perspective of this project. There are several characterizations of the ZPD (Vygotsky, 1987); this report will focus on the scaffolding; where a learner is in their ZPD if they can complete a problem with assistance they could not complete independently. This characterization of ZPD dovetails with the second purpose of Black & Wiliam's framework (2009): formative assessment is used to engineer effective classroom discussions; where scaffolding may be given to a group of students in an efficient manner.

Participants were recruited from two introductory calculus courses taught using the approximation framework. This framework is built upon developing systematic reasoning about conceptually accessible approximations and error analyses but mirroring the rigorous structure of formal limit definitions and arguments (Oehrtman, 2008, 2009). This study focused on the three multi-week labs developing the most central topics in the course: limits, derivatives, and definite integrals. Each approximation lab consists of 20 questions designed to help students understand their context in terms of approximating a limit (Figure 1).

	Contextual	Graphical	Algebraic	Numerical
Unknown Value				
Approximation				
Error				
Error Bound				
Desired Accuracy				

Figure 1. Approximation framework

The courses were taught at the same time and on the same schedule by two equally experienced instructors. All of the lab questions were scored dichotomously so the inter-rater reliability of the lab write-ups was perfect, and the final exams were co-graded by the instructors. The content validity of the assessments was checked by the course coordinator and an additional expert on the approximation framework. All assessments had reliabilities within acceptable levels (Gall, Gall & Borg, 2007): the limit, derivative, and definite integral labs had KR-20 values of 0.83, 0.72, and 0.78 respectively; the final exam had a Cronbach Alpha of .68.

In addition to participants' lab write-ups, the final exam and demographic information was collected from each participant. There were no significant differences between the sporadic and non-participation groups on all but one of the demographic or grade predictive variables tested ($p > 0.25$)[1]. Female students were significantly more likely to be regular participants in formative assessment ($p = .03$). Since asynchronous formative assessment, like the ones used in this study, require a greater level of organization and engagement, these assignments tend to slightly favor female students (DiPrete, 2013). Despite the selection bias inherent in the participation groups, there was no evidence at the start of the semester to suggest that students at different participation levels would perform differently in the course.

There were 66 students that consented to participate in the study; 13 of the students were removed from the sample because they had prior exposure to the labs that could confound the results. Of the 53 students that were new to the approximation framework labs, only seven had no prior exposure to limit concepts in a prior course, and 27 of the students had AP Calculus in high school. There were 14 students classified as sporadic participants in formative assessment and 16 students classified as non-participants; the remaining 23 students participated regularly in the formative assessments (Table 1).

Table 1
Gender distribution of the participation groups

Group	Male	Female
Regular	8	15
Sporadic	9	5
Non-Participant	11	5

Each assessment was analyzed using one way ANOVA; Tukey-Kramer tests were used when there was a significant difference found on the ANOVA. To investigate the effects of the formative assessment-based class discussion on achievement on the summative

[1] Age, race, native language, and year in school showed no significant difference on a Chi-Square test. GPA, Math GPA, ACT Math Score, Math GPA, the pre-calculus skills test administered the second day of class, and time elapsed since the previous mathematics course showed no significant differences using Mood's median test.

assessments, item discriminations were calculated on all items discussed by the instructors in class.

Results

The limit lab asked students to approximate the location of a removable discontinuity where there were no obvious algebraic manipulations that would allow the discontinuity to be calculated exactly. Much of the lab depended on familiarity with function concepts. Given that there was only one formative assessment based discussion and there were no significant differences between the participation levels in any prior knowledge measure available, it is not surprising that the ANOVA found no significant differences in group achievement on the lab write-up (Table 2); the context of the lab was equally familiar to all students and there were not enough instructional interventions to make a difference.

Table 2
Results of the limits lab achievement scores

	SS	Df	MS	F	p
Between	50.637	2	25.32	1.702	0.193
Within	758.693	51	14.876		
Total	809.33				

The derivative lab asked students to approximate an instantaneous rate of change at a given point using the slopes of secant lines. During this lab, there were three formative assessment-based discussions, and students were given an opportunity to turn in a draft of their write-up for formative feedback. The ANOVA results shown in Table 3 revealed a significant difference in achievement on the lab write-up between the three participation levels. According to the Tukey-Kramer test, the significant difference in achievement was between the regular participation group and the sporadic and non-participation groups. The sporadic and non-participation groups' achievement was not significantly different from each other. The difference between the regular participation group and the others is that the regular participants consistently used slopes as their approximations; the sporadic and non-participation groups completed this lab using y values as approximations more often than the regular participation group.

Table 3
Results of the derivatives lab achievement scores

	SS	Df	MS	F	p
Between	303.21	2	151.61	13.96	0.000
Within	553.96	51	10.862		
Total	857.12	53			

The definite integration lab asked students to model a given quantity with a definite integral and then approximate their quantity with Reimann sums. There were two formative assessment-based discussions during this lab; these discussions focused on summation notation and assistance with the technology required to calculate large Riemann sums. The ANOVA results shown in Table 4 revealed a significant difference in achievement between the three participation levels. The Tukey-Kramer test revealed that all three groups were significantly different; the regular participant group outperformed the sporadic participant group who in turn outperformed the non-participant group. The cumulative common final exam ANOVA and the corresponding Tukey-Kramer test had similar results to the definite integral lab; all three groups had significantly different levels of achievement from each other, and were in the same order (Table 5).

Table 4

Results of the definite integral lab achievement scores

	SS	Df	MS	F	p
Between	1108.39	2	550.69	12.252	0.000
Within	2292.31	51	44.95		
Total	3393.70	53			

Table 5

Results of the final exam achievement scores

	SS	Df	MS	F	p
Between	19881.53	2	9940.77	20.968	0.000
Within	24179.40	51	474.09		
Total	44059.92	53			

Discussion

While these results indicated that there were measurable achievement differences between the three participation groups, the more interesting, and more difficult, question is why these differences exist. There appear to be two plausible explanations based on the available data. The first is that the formative assessment-based instruction was more effective for the students participating in the formative assessments. After conducting an item discrimination analysis on the lab write-ups, 48% of the items that were discussed in class showed no significant discrimination between participation levels. While this result suggests that all participation levels benefit on some level from the in-class discussions, the non-participant group had fewer non-discriminating items on every assignment; on the final lab only one question failed to discriminate between the regular participant group and the non-participant group. There were three non-discriminating items between the sporadic and non-participant groups on the final lab.

The other plausible explanation is these achievement patterns are indicative of a lurking variable. The questions we will pose to those attending our talk will explore this explanation: (1) Can any further analysis be done with the current data? (2) Are there any constructs other than attribution that could account for the differences in achievement?

References

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, *21*(1), 5-31.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (in press). Meta-Analytic Methodology and Conclusions about the Efficacy of Formative Assessment. To appear in *Educational Measurement: Issues and Practices*.

DiPrete, T. A. (2013). *The Rise of Women: The Growing Gender Gap in Education and what it Means for American Schools*. Russell Sage Foundation.

Gall, M.D., Gall. J.P., & Borg, W.R. (2007). *Educational research* (8th ed.). New York: Pearson.

Karpinski, A. C., & D’Agostino, J. V. (2012). The role of formative assessment in student achievement. *International Guide to Student Achievement*, 202.

Oehrtman, M. (2008). Layers of Abstraction: Theory and Design for the Instruction of Limit Concepts. In M. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Practice in Undergraduate Mathematics* (pp. 65-79). Washington, DC: MAA.

Oehrtman, M. (2009). Collapsing dimensions, physical limitation, and other student metaphors for limit concepts. *Journal for Research in Mathematics Education, 40*, 396–426.

Vygotsky, L. (1987). The development of scientific concepts in childhood. In R.W. Rieber and A.S. Carton (Eds.) *The collected works of L.S. Vygotsky, Volume 1: Problems of general psychology* (pp. 167-241). New York, NY: Plenium Press.

55

STUDENTS' EXAMPLE USAGE IN THE DOMAIN OF FUNCTIONS

Muhammed F. Dogan
University of Wisconsin - Madison

Abstract: Mathematicians use examples strategically while working on mathematical conjectures, and this strategic usage helps them gain a lot of insight about mathematical phenomenon. However, students do not always have the same strategic example usage; instead, they tend to over rely on examples without understanding of example based reasoning. This study examines college algebra students' responses on a written assessment in the function domain and discusses students' example spaces. The results reveal that students have very limited example space in the function domain that affects their strategic example usage. Student example usage was very limited to conventional example spaces that they learned during instruction or from their textbook. This study suggests that having conventional example spaces does not guarantee that students can use examples strategically which can help them better understand the mathematical conjectures.

Key Words: Algebraic reasoning, Function, Example space, Exemplification

Introduction

Even though the mathematics education literature has indicated that proof activities are a crucial element in mathematical thinking and activity, mathematics education research has shown that students fail to understand the nature of evidence and justification in proving activities in mathematics (e.g., Harel & Sowder, 1998; Knuth, Choppin, & Bieda, 2009). Researchers give a number of reasons for students' struggles with proving, including students not understanding the importance of proof, not being able to articulate the mathematical knowledge that they have, and not feeling comfortable with mathematical activities. One additional reason posited in the mathematics education literature is based on students' treatment of examples in proving activities (e.g., Stylianides & Stylianides, 2009; Healy & Hoyles, 2000; Porteous, 1990), as students tend to be overly reliant on examples and generally believe that they have proved a mathematical statement by using a few examples that satisfy it. While acknowledging the limitations of examples in proving activities, researchers have also shown that examples can be very crucial in helping students generate proof (e.g. Ellis et al., 2012; Lockwood et al., 2012; Alcock & Inglis, 2008).

The purpose of this study is to better understand the role that examples play in the exploration and justification of mathematical conjectures, especially in domain of function in college level mathematics. Functions are central to numerous mathematical concepts, and they play a vital role in advanced mathematical thinking. In the study described in this proposal, I examine how students use examples while working with conjectures involving odd, even, and one-to-one functions. The research question is: How do students use examples, and what is the nature of their thinking about the examples they use to develop, explore, understand, and prove mathematical conjectures about functions?

Literature Review and Theoretical Background

It is well known that examples play an important role in mathematicians' thinking as they engage with mathematical activities. Watson and Mason (2002, 2005) see examples as the heart of mathematical learning at all levels. However, the idea of examples is very broad, and it is unclear how students use examples to develop an understanding of a mathematical concept or to explore and prove a mathematical conjecture. There are a number of studies that focus on the limitation of examples use. Most of those studies claim the same idea, which is that students overly rely on examples and tend to convince themselves that something is true by using a few examples while engaging in proving activities (e.g. Harel & Sowder, 1998; Bell, 1976; Balacheff, 1987; Stylianides & Stylianides, 2009). Therefore, students do not see a need to develop a proof that goes beyond showing examples; they think they have sufficiently proved a conjecture and do not understand the limitations of using examples as proof.

In the literature, the role of examples in proving activities has recently gained a lot of attention. Researchers have claimed that examples are not solely a barrier in proving activities, but instead can be a crucial component of understanding and developing a mathematical proof when examples are used strategically (e.g. Ellis et al., 2012; Lockwood et al., 2012;; Weber & Mejia-Ramos, 2011; Alcock & Inglis, 2008; Buchbinder & Zaslavsky, 2011; Antonini, 2006; Dahlberg & Housman, 1997; Iannone et al., 2011). It is important to note that these studies deemphasize the limitations of the use examples in proof, but rather claim that examples also can be very powerful, especially when developing reasoning about proof. Overall, the literature suggests that examples are an important part of mathematics in developing conceptual understanding of a mathematical concept. I have designed my study with this literature in mind, drawing upon the fact that example-based reasoning can be very powerful in terms of learning, understanding, and making sense of mathematical conjectures.

Theoretically, I frame the study in terms of the notion of an example space, which is based upon the notions of concept definition and concept image by Tall and Vinner (1981) and Vinner (1993). Tall and Vinner (1981) and Vinner (1993) define concept definition as the words and symbols used to specify a concept, and concept image as “the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes” (p. 2). Students cannot always reason and justify mathematics by using a concept definition, so they need a concept image to reason and justify about mathematical properties. Tall and Vinner (1981) describe a learning gap between students’ concepts definitions and concepts images, and point to the importance of using examples to close this gap. Thus, the ideas of concept image and concept definition are closely related to use of examples. Additionally, Watson and Mason (2005) discuss examples in a broader sense by introducing the idea of *exemplification* in mathematics and the idea of *example space*. They use the term exemplification to “describe any situation in which something specific is offered to represent a general class with which the learner is to become familiar—a particular case of a generality” (p. 4). Exemplification itself is a pedagogical tool for mathematical learning, but it is also worth mentioning that it is crucial to allow students to generate their own examples and exemplify the concept for themselves since example creation and generation is individual and situational (meaning an example may make sense in one situation but not in a different situation).

Methods

The data presented here consists of undergraduate students’ responses to a subset of items from a written assessment targeted to understand how students’ use mathematical examples in tasks involving functions. The assessment was designed to have students interpret mathematical definitions, use the definitions, generate examples, demonstrate the role of examples in their work, and produce convincing mathematical arguments and proof. The results shared in this paper focus on two main items from this study that asked students' to generate justifications regarding the truth of mathematical conjectures. 105 college students at a large mid-western university responded to the questions on the written assessment. All of the participants were enrolled in a "College Algebra" course and all had taken a "Pre-calculus Math" course the previous semester. Data were collected in seven different College Algebra classrooms. The researcher collected data by attending the last 30 minutes of each of the seven classes. Table 1 shows the conjectures that students were given on the assessment.

Table 1 - Conjectures

Even and Odd Function Tasks	**Let *f* and *g* be functions with domain all real numbers. Answer the following:** a. Suppose f and g are even functions. Is $f - g$ is an even function, an odd function, or neither? **Please show your work and justify your answer.** b. Suppose f is an even function and g is an odd function. Is the sum $f + g$ an even function, an odd function, or neither? **Please show your work and justify your answer.** c. Suppose f is an odd function and g is an even function. Is the composition $f \circ g$ even, odd, or neither? **Please show your work and justify your answer.**
One-to-One Function Tasks	**Let *f* and *g* be functions with domain all real numbers. Please answer the following:** a. Suppose f and g are one-to-one functions. Is $f + g$ also one-to-one? **Please show your work and justify your answer.** b. Suppose f is a one-to-one function and c is a nonzero real number. Is cf also one-to-one? **Please show your work and justify your answer.**

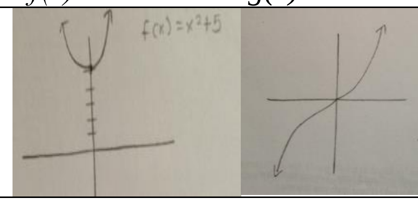
The tasks were presented in three different groups depending on the types of examples given in the task, and the students were randomly assigned to groups. All three of the groups were given definitions for even, odd, and one-to-one functions, as well as the same five mathematical conjectures (Table 1) to evaluate., The groups varied in terms of which examples were provided with the conjectures and definitions. Group I was provided with an algebraic example and a graphical example for the three kinds of function (even, odd and one-to-one). Group II was given the definition of the functions and with the items, and they were explicitly asked to provide an example for each kind of function and were not given examples. Group III was given the definitions of functions and the items, but was not provided with or asked to generate any examples. At the time when this data was collected, the participants were studying polynomial functions. The instructors of the course claimed that most of their students could successfully solve the items in the given time because they had studied even, odd, and one-to-one functions within the previous two weeks. The instructors assumed their students knew all of these concepts since they had worked on them extensively.

Analysis of the written assessment consisted of examining the data for emergent themes and categories. Open coding (Strauss & Corbin, 1998) was used to develop categories classifying the different types and uses of examples that students' gave to develop, explore, understand, and justify mathematical conjectures.

Results and Discussion

As a result of the open coding, three main categories of example types were identified: graphical examples, algebraic examples, and non-examples. Table 1 shows definition of each example type and examples of it. Themes also emerged from the analysis, and these are discussed below.

Table 2: Example Types

Example Type	**Definition**	**Example**
Algebraic Examples	Use algebraic symbols to represent examples	“$f(x) = x^2$ and $g(x) = x^3$” “$f(x) = x^2 + 4$ and $g(x) = x^4$”
Graphical Examples	Use graphs to represent examples	
Non-examples (irrelevant examples)	Examples that are not related to the concepts, which means they are not right examples of functions.	F(x)=2; as example of one-to-one function: $f(x) = x^2$

As seen in Table 2, the participants generated mostly non-example which meant the examples that they used were not related to the concept they were asked about. Algebraic examples were the next most common type of examples used.

Example Type	**Group I**	**Group II**	**Group III**	**Total**
Algebraic Examples	71	37	33	141
Graphical Examples	18	12	19	49
Non-examples	111	99	120	330

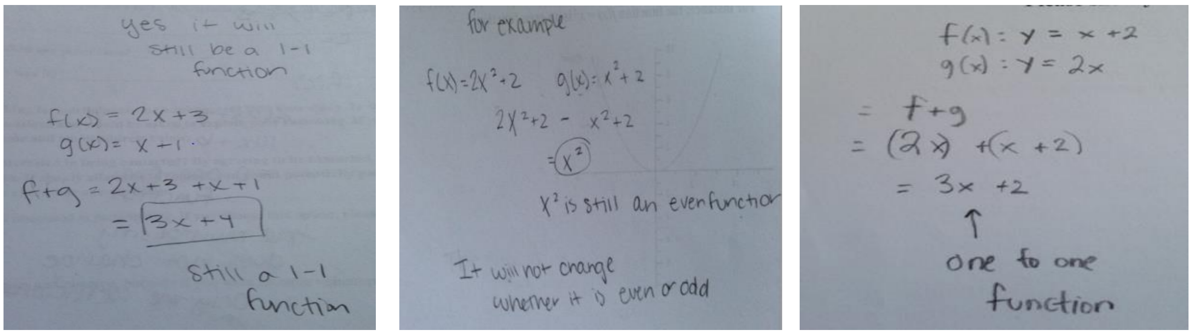
Table 3: Counts of Groups’ Example Types

Group I. Students in Group I were provided with an algebraic example of an even function (x^2), an odd function (x^3), a one-to-one function (x^3), and a non-one-to-one function (x^2), as well as with a graphical representation of the algebraic functions. The examples given were from their textbook, so since those examples were familiar to them from having seen them in the textbook and during instruction, the examples are defined as examples from conventional example space (Watson & Mason, 2002; 2005). In general, students stuck with the provided examples and did not generate other examples to work with. Students generated functions that were very similar to x^2 and x^3 functions which were given to them, such as x^4, x^5, and x^6. This suggests that students’ example spaces were limited which would affect the examples they would choose and may also affect how they use those examples.

Group II. Students in this group were asked to provide an example for each definition. The students generated examples that were algebraic (49 cases), graphical (45), non-examples (6) and other-examples (4). Interestingly, these students generated similar examples to the ones provided to Group I, which suggests that this group had conventional example spaces based on the textbooks or instruction used in their class.

Group III. Students in this group were just given the definitions and the items without any prompt or examples. Almost half of these students exemplified the items (that is, they used examples to make sense of the conjecture by generating their own examples (Watson & Mason, 2002; 2005)) while the other half did not use examples or definitions. Interestingly, all of the students who used examples to work with the items had similar types of examples as the ones given for Group I and Group II. This suggests that their example spaces were limited to conventional example spaces as well.

In addition to these results, none of the students in this group attempted to make a successful general argument by proving the conjectures, but they demonstrated empirical proof scheme (Harel & Sowder, 1998) by trying a few examples then generalizing from those examples. Figure 1a-c shows responses to questions are some examples of their empirical proof scheme:



(a) (b) (c)

Response (a) and (c) show that students used an example for each function and showed that the conjecture is true for that case, but they did not have a general argumentation for it. Indeed, they responses were true for those cases, but the conjecture was a false conjecture so that their responses were not correct. Similarly, response (b) shows that a student used two similar examples for even functions and made a general argumentation by testing only one case. Her response was true, but she did not provide an explanation for it, and this was very common.

In addition to the result shared above, there are some important results that this study revealed, which suggest that students do not have robust understandings of functions. The first one is that there was more non-examples type than all of other type of examples about the concept of functions. Indeed, students mostly used examples that are not representation of functions, saw functions just as an "input and output," and were not able to use the basic definition of the function to prove the conjectures or to generate examples. Additionally, students think about functions as numbers and treat numbers as a general case of functions, which suggests that they do not have understanding of functional thinking. For example, the following responses suggest that students think of even and odd functions as they do even and odd numbers: *"Odd function, because when subtracting an odd from an even you will always get odd."*

Or, they see even and odd functions as opposites as they do negative and positive numbers: "*Neither, because you are adding functions that are opposite, one being positive and one being negative"*

In addition, many students mention that they needed an actual function to work with:

"What is the original function to test with?"

"I don't know what f(x) or g(x) is so I can't find a solution to the problem."

Overall, these results show that there was no difference between the groups in terms of example type usage. All three groups used similar kinds of examples types, which suggest that their example spaces were similar. A possible explanation for this is that their example spaces were based on the textbook or the instruction used in their class. This finding highlights the importance of example space. If students don't have access to a robust example space, they are not going to be able to have access to examples and to use examples productively. Indeed, Watson and Mason (2002, 2005) claim that "the learners' potential example space is likely to be a subset of what is taken to be conventional by mathematicians and textbooks" (p. 59). One possible explanation for these results is that to accommodate a new mathematical definition, students need to generate something new from their example spaces in order to expand their example spaces and to move to an abstract level of thinking, which did not happen in this study because of their limited example spaces.

In the first part of the result section, I discussed how students' example spaces were limited to conventional examples, which is indicative of a potential limitation of students' concept image of functions. Students did not understand the properties of the examples that they used since the examples they used were based on their conventional example spaces. Watson and Mason (2002, 2005) claim that "conventional example spaces are generally understood by mathematicians and as displayed in textbooks, into which the teacher hopes to induct his or her students" (p. 76). However, as the data shows, these hopes were not achieved in my study. The main reason for this, I believe, is that students' concept definition and concept images about functions were not well developed, and students were not able to articulate their understanding, which suggests that they were not able to accommodate the knowledge that they encountered. Finally, not surprisingly, students tried to use empirical proof schemes (Harel and Sowder, 1998) often, but, due to their limited functional thinking and problems with algebraic manipulation, they could not achieve empirical proof scheme either, in general.

Conclusion

The purpose of this study was to investigate how students use examples while exploring mathematical conjectures about functions. The results showed that student example usage was very limited to conventional example spaces, which were likely based on the instruction and textbook used in their college algebra course. This suggests that having conventional example spaces does not guarantee that students can use examples strategically. Since students did not have a robust understanding of functions, they could not use examples to make sense of the conjectures and to generate proofs. In addition, students did not generate a variety of examples when they were asked to do so, or they did not know how to use examples when they were provided with examples. They could not expand their example spaces in terms of giving examples because the domain of function was so far outside of their experiences that they could not articulate their examples. Therefore, generating examples is not easy for students. Instead of seeing examples as obstacles in proving activities, we should think of ways for how to help student generate and use examples.

References

Alcock, L. J. & Inglis, M, (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69*, 111-129.

Antonini, S. (2003). Non-examples and proof by contradiction. In N. Pateman, B.J. Dougherty, & J. Zilliox (Eds.), Proceedings of the 27th Annual Meeting of the International Group for Psychology in Mathematics Education (Vol. 2, pp. 49 – 55). Honolulu: University of Hawaii.

Balacheff, N. (1987). Treatment of refutations: Aspects of the complexity of a constructivist approach to mathematics learning. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 89–110). The Netherlands: Kluwer Academic Publishers.

Bell, A. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics, 7*, 23–40.

Buchbinder, O., & Zaslavsky, O. (2011). Is this a coincidence? The role of examples in creating a need for proof. *ZDM - Zentralblatt fuer Didaktik der Mathematik*, *43*(2), 269-281.

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics*, *33*(3), 283-299.

Ellis, A. B., Lockwood, E., Williams, C., Dogan, M. F., & Knuth, E. (2012). Middle school students' example use in conjecture exploration and justification. In *Proceedings of the 34th Annual Meeting of the North American Chapter of the Psychology of Mathematics Education*.

Healy, L. & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, *31*(4), 396–428.

Harel, G., & Sowder, L. (1998). Students' proof schemes. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Research on Collegiate Mathematics Education* (Vol. III, pp. 234–283). Providence, RI: American Mathematical Society.

Iannone, P., Inglis, M., Mejia-Ramos, J. P., Simpson, A., & Weber, K. (2011). Does generating examples aid proof production? *Educational Studies in Mathematics, 77,* 1–14.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). New York, NY: Routledge.

Lockwood, E., Ellis, A.B., Dogan, M.F., Williams, C., & Knuth, E. (2012). A framework for mathematicians' example-related activity when exploring and proving mathematical conjectures. In L.R. Van Zoest, J.J. Lo, & J.L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 151 – 158). Kalamazoo, MI: Western Michigan University.

Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics, 21,* 589–598.

Strauss, A. & Corbin, J. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, *40*, 314–352.

Tall, D.O. and Vinner, S.: 1981, 'Concept image and concept definition in mathematics, with particular reference to limits and continuity', Educational Studies in Mathematics 12, 151-169.

Vinner, S. (1983). Concept definition, concept image and the notion of function. International Journal of Mathematics Education in Science and Technology, 14, 293-305.

Watson, A., & Mason, J. (2002). Extending example spaces as a learning/teaching strategy in mathematics. In *PME CONFERENCE* (Vol. 4, pp. 4-377).

Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Routledge.

Weber, K., & Mejia-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics*, *76*(3), 329-344.

Formal Logic in Early Undergraduate Mathematics: A Cycle

Morgan Dominy
Virginia Tech

Topic: Examining the understanding of formal logic by early undergraduates leads to the examination of a cycle within the broader scope of math education. Some early undergraduate students will eventually become pre-service teachers. The pre-service teachers will in turn become primary and secondary educators. Finally over time, some of their students will become undergraduates leading to a cycle. This poster proposes the idea that any change of curricula in undergraduate mathematics should have a ripple effect on the overall understanding of formal logic by students of all levels of education over a period of time as the cycle flows.

Background literature: Several papers serve as motivation for the idea. A. and G. Stylianides have developed a course to improve the understanding of formal logic by pre-service teachers (Stylianides & Stylianides, 2009). A. Stylianides (2007) deals with the concept of proof in the setting of primary education by allowing theorems in mathematics to be taken as axioms in a classroom setting. For example basic arithmetic operations and their properties are proved in the theory using the foundational axioms, but in a classroom settings student intuition is built upon by allowing the properties themselves to be taken as axioms. D. Thomson, *et al* conducted a study of published high school math texts and found that less than 6% of the exercises involved proof-related reasoning (Thomson, *et al*, 2012). M. Inglis and L. Alcock made the use of modern technology to objectively measure the "warrant-seeking" behavior of both mathematicians and undergraduates when reading proofs. (Inglis & Alcock, 2012).

Research potential: Just as Inglis and Alcock have measured student understanding of proofs, it should be possible to devise similar methods to quantify and measure student understanding of formal logic on all educational levels. If done over a long period of time, the effects of the implementation of ideas such as the course designed by the Stylianides can be measured. Additionally by comparing results from different educational levels, it should be possible to identify the level of education that requires the most improvement leading to a sort of triage.

Current focus: In many undergraduate mathematics curricula, a course in formal logic is not a prerequisite for earlier courses such as calculus, differential equations, and linear algebra. However all three of these subjects rely upon student understanding of formal logic when solving non-routine problems related to concepts such as limit, infinite series, and linear independence. A possible avenue of quantitative research in this area would involve measuring student responses to non-routine problems from the aforementioned courses and comparing the responses of the students who had taken a formal logic course or are currently enrolled in one to those who had not. This may lead to the offering of equivalent courses with a greater instructional focus on the logical concepts covered in the course. Any such change in curricula will cause a far reaching effect over time that can also be studied. The purpose of this poster is to start a dialogue about this idea among math educators and obtain valuable insight about what direction to take with a research program.

References

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43(4), 358-390.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 289-321.

Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics*, 72(2), 237-253.

Thompson, D. R., Senk, S. L., & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education*, 43(3), 253-295.

CALCULUS STUDENTS' UNDERSTANDING OF UNITS

Allison Dorko
Oregon State University
dorkoa@onid.orst.edu

Natasha Speer
University of Maine
speer@math.umaine.edu

Units of measure are critical in many scientific fields. While instructors often note that students struggle with units, little research has been conducted about the nature and extent of these difficulties or why they exist. This study investigated calculus students' unit use in area and volume computations. Seventy-three percent of students gave incorrect units for at least one task. The most common error was the misappropriation of length units in area and volume computations. Analyses of interview data indicate that some students think that the unit of the area or volume computation should be the same as the unit specified in the task statement. Findings also suggest that some students have difficulties correctly indicating the units for computations that involve the quantity π. In addition, findings suggest that calculus students' difficulties with units are linked to their difficulties with understanding area and volume as arrays.

Key words: undergraduate students, units, area, volume, arrays

Background and Research Question

Units of measurement are important in many disciplines including those in science, technology, engineering and mathematics (STEM). Many of our future scientists and engineers get their foundational understanding of the key ideas of measuring quantities (e.g., rates, rates of change) in an undergraduate calculus course. Success in any STEM discipline requires both the knowledge needed to obtain quantitative answers and the knowledge of how those quantities relate to the physical world, including the units in which such quantities are measured. Findings from research indicate that many of these sorts 'applications' of mathematics, such as optimization (min-max problems), related rates, and volumes of solids of revolution are difficult for students (Engelke, 2008; Martin, 2000; Orton, 1983). It is not known how students' knowledge of spatial measure and its units interact with the calculus necessary in these problems, and clarifying this is potentially useful for bettering instruction. Other researchers have noted how useful units are in making sense of a problem and knowing what quantities to combine, yet students rarely have these sorts of skills (Redish, 1997; Rowland & Jovanoski, 2004; Rowland, 2006; Saitta, Gittings, & Geiger, 2011). Investigating calculus students' understanding of units has implications for calculus instruction and possibly for students' unit understanding in other fields.

The purpose of this study was to investigate differential calculus students' success with and understanding of the units for area and volume computational problems. The few studies that exist about undergraduate students' understanding of units provide evidence that units pose difficulties. Chemistry education researchers have documented that dimensional analysis is difficult for students, even though unit conversions are often taught in a variety of subjects over a variety of grade levels (Saitta, Gittings, & Geiger, 2011). In a study about students' understanding of the units in differential equations, researchers found that many students did not understand that both sides of the equation must have the same units (Rowland, 2006; Rowland & Jovanoski, 2004). Students also struggled with determining the units of a proportionality constant and explaining the meaning of the terms in the equation (Rowland, 2006; Rowland & Jovanoski, 2004). Redish (1997) found that physics students may interpret the symbols in an equation as pure numbers rather than as standing for physical quantities.

There is a broader base of literature about elementary school students' understanding of units and spatial measure. These students frequently misappropriate units of length for area calculations and do not see the need for a unit of cover (Iszák, 2003; Fuys, Geddes, &Tischler, 1998; Lehrer, Jenkins, & Osana,1998). Elementary school students also have difficulty using the unit structure of an array of cubes to determine the volume of a rectangular prism (Battista & Clements, 1998; Curry & Outhred, 2005). These findings indicate that elementary school students struggle with dimensionality, which may result in difficulty determining the correct units of measure.

Given the importance of units and their use in computations modeling physical situations, it seems worthwhile to investigate undergraduate students' understanding of units and to see if the issues present in elementary school students persist into students' university years. We surveyed and interviewed 198 differential calculus students to find out the following:

- What percent of differential calculus students write correct units for area and volume computational tasks? What is the thinking and reasoning of these students?
- What percent of differential calculus students write incorrect units for area and volume computational tasks? What is the thinking and reasoning of these students?

In the next section, we describe the tasks we used to investigate these questions and our method of analysis.

Method

We collected data from differential calculus students at a large northeastern university. First, we had 198 such students complete the following area and volume computation tasks:

- Area of a rectangle (12 cm x 4 cm)
- Area of a circle (r = 5 in)
- Volume of a rectangular prism (5 cm x 4 cm x 10 cm)
- Volume of a cylinder (r = 3 in, h = 8 in)
- Volume of a right triangular prism (prism h = 8 ft; triangle base = 4 ft, triangle h = 3 ft)

We then analyzed these data and interviewed students whose survey responses fell into the categories identified in the written data. The interview tasks were the same as the written tasks. The strength of our survey-then-interview method was that we are able to report quantitative data about students' success rates on the problems as well as qualitative data regarding the sorts of thinking and reasoning behind correct and incorrect answers.

We used a Grounded Theory inspired approach to data analysis (Glaser & Strauss, 1967). Specifically, we looked for commonalities across responses and developed codes and categories based on these emergent themes. We chose to look at the unit only, and not the magnitude, in each response. We assigned codes to each response, not to each student; that is, we looked at each answer individually. The first round of coding was for correct units, incorrect units, and no units. On the area tasks, we marked as 'correct unit' any squared unit (e.g. cm^2, in^2 , $units^2$) and any cubed unit (e.g. cm^3 , ft^3 , $units^3$) on the volume tasks. Any non-square unit on the area task or non-cubic unit on a volume task was marked as 'incorrect unit' (e.g., cm, in, ft^4). Any magnitude without a unit was marked as 'no unit.' This coding allowed us to answer the first part of the two research questions regarding students' success with area and volume tasks. We found percentages of correct/incorrect/no units (Table 1) for each task, as well as how many students had correct units for all of the tasks.

Following this first layer of coding, we looked for patterns in the 'incorrect unit' and 'no unit' categories. We noticed two themes: many students gave length units for area/volume computations, and the responses to the circle and cylinder tasks had a higher proportion of 'no units' than the other tasks. This led to two more rounds of coding. The first was to look at the 'incorrect unit' responses and see how many of those were length units such as 'cm' or 'in' (Table 2). The second was to investigate the circle and cylinder tasks. We had noticed that while many of these responses did not have a unit, they did include the symbol π. In this coding, we coded how many students had no units for either the circle task, the cylinder task, or both. Based on interview data (discussed later), we termed these categories 'π with no unit,' (e.g., 72π, 25π) and 'answer with unit' (e.g., 72π in, 25π in^2). We also looked at the units students used with the other tasks.

We then turned to our interview data to learn the thinking and reasoning behind students' answers. We coded the area/volume computation in the interview as 'correct unit,' 'incorrect unit,' or 'no unit,' and then used Grounded Theory to look for patterns in the collection of interview excerpts that fit each category. For example, we identified interviewees who had all of the units correct and looked for patterns in their reasoning. A number of students talked about arrays and/or dimensionality, so we paid attention to the appearance of those words and ideas as we analyzed data. As a second example, students with incorrect or no units for the circle and cylinder problems often talked about the symbol π as troublesome. Details of how transcript excerpts were coded are presented in conjunction with the results. In the next section, we present our quantitative results and the findings from the interview data about the thinking and reasoning behind students' answers.

Results & Discussion

Table 1 shows the results of the first round of coding, which was the broadest level to determine what percentage of students used correct units, incorrect units, and no units on the tasks.

Table 1. Correct, Incorrect, and No Units

Task	Area of rectangle	Area of circle	Volume of rectangular prism	Volume of cylinder	Volume of triangular prism
n total	197	197	197	197	128
n responses	195	195	196	182	107
Correct unit % of responses	159 81.54	113 57.95	165 84.18	125 68.68	77 71.96
Incorrect Unit % of responses	26 13.33	27 13.85	23 11.73	23 12.64	19 17.76
No unit % of responses	10 5.13	55 28.21	8 4.08	34 18.68	11 10.28

The percentages in the table are the percent of the total responses (not all students answered all questions). For instance, 84.18% of students had the correct unit for the volume of the rectangular prism task. We also looked at how many students had correct units for all of the tasks, and found that of 169 students who answered all of the tasks, 45 (26.6%) had correct units with all of their responses.

Interview data helped us understand how students thought about the units of the problems. We found that interviewees who had correct units talked about at least one of the following: dimensionality, arrays, and rules of exponents. For instance, Steven and Isaac had correct units for all of the problems. They explained how they decided what unit was correct as follows:

> **Steven:** Area is a square [unit]. Every time we multiply one dimension by the next, we're multiplying centimeters by themselves. We change from a linear to an area, then area to volume. It's recognizable that volume is cubic, as opposed to area, which is squared. [The volume of the rectangular prism] is the length times the width of the base times the height. It's like the area formula only now we have a vertical dimension so we have 200 little centimeter boxes inside.
>
> **Isaac:** If we think about this [5 cm x 10 cm face] in terms of an area – you have 50 boxes [draws a 5 x 10 array of squares on the face]. You know that you have four of these, fifty times four. So you can think of it as having four layers of this area because the difference between volume and area is just adding another dimension … so we can think of it as four sheets of 50 squares.

Steven thought about dimensions and rules of exponents as he identified a correct unit. Isaac, who determined that the volume of the box was 200 cm^3, thought about arrays, and used the idea of sheets of squares (we suspect he meant 'sheets of cubes') to reason that the unit should be cubic centimeters. These students' thinking is representative of others who had all the units correct. Therefore, we can conclude that ideas of dimensionality, arrays, and rules of exponents are helpful for students to understand the units of a spatial computation.

Table 2 shows the results of coding for length units.

Table 2. Misappropriation of Length Units

	Area of rectangle	Area of circle	Volume of rectangular prism	Volume of cylinder	Volume of triangular prism
n total	197	197	197	197	128
n responses	195	195	196	182	107
Length Units % responses	29 14.87	25 12.82	18 9.14	13 7.14	16 15.00

One finding about elementary school students' understanding of units is that some elementary school students believe length to be an adequate measure of area (Lehrer, 2003). As evident from the table above, this belief may persist in some undergraduates. Though only two students used length units on all of the tasks, a number of students used length units on one or more of the tasks (as detailed in Table 2). Interviewees who used length units for all the tasks told us that the unit of the computation was [to paraphrase] 'the same as the unit you're told in the problem.' Rae, Alex, and Jolie were three such students; they wrote "cm" with their answers. Here, they respond to the interviewer asking them to "tell me about the centimeters."

> **Rae:** That's the thing that goes along with the sides. Whatever the side is in, that's what the answer is in.
>
> **Alex:** We were given a unit. The numbers are centimeters.

While students who had correct units thought about dimensionality, rules of exponents, and arrays, there is no indication from Rae and Alex that they were thinking about these things. This supports our conclusion of the importance of understanding dimensionality, rules of exponents, and arrays in understanding the units of spatial computations. We provide one final example to support this. Jolie wrote length units for all of the tasks except the triangular prism, in which she wrote '60'. The interviewer directed her back to the area of the rectangle task:

> **Interviewer:** On this rectangle question, you wrote 48 centimeters. Does this question have anything like that [points to the 'cm']?
>
> **Jolie:** Well, it's feet, but I don't know if it's cubed because it's a triangle. I know it wouldn't be just feet.
>
> Interviewer: Do you remember volumes of other things being cubed when you've done volume problems before?
>
> **Jolie:** I don't remember.

Jolie knew the measurement was not just feet, but was not sure whether or not it was cubed. Her hesitation may have been due to a perceived mismatch of a triangular shape and cubic units. That she was unsure about the units suggests that she did not understand the connection between units and dimensionality. This seemed to be the case for many of the interviewees who struggled with units.

Coding the circle and cylinder tasks for "π with no unit" revealed that 55 of 195 students (28.21%) had an answer like 25π for the circle task and 34 of 182 students (18.68%) had an answer like 72π for the cylinder task (the magnitudes of the answers may have differed). This is interesting because the percentages are much higher than the 'no unit' percentages for the other tasks (see Table 1). We found 10 students who had π and no other units for both of the tasks (these 10 students are a subset of the 55 and 34). Eight of these ten students had given units for all of the other tasks, and the other two students did not give units with any of the tasks. That eight of these students included units with the other tasks seems a compelling piece of evidence that something about circle and cylinder area/volume computations causes students to have issues with the units. While a possible alternative explanation is that the students 'forgot' the units or were being 'careless,' interview data leads us to believe otherwise. For instance, Amy and Bob said that they "forgot" the unit because of the π involved:

> **Amy:** I probably didn't even think of it because I was using pi, so I left pi in it and I didn't think to label it. But I labeled all the rest of them. That's really weird. Well I know pi is an actual value, but I guess I would … I don't know. It probably just slipped my mind because I was using pi to represent a number rather than saying 3.14 and I probably just forgot to put a label on it. I probably have a tendency to do that with circles because you really only use pi with circles and it kind of doesn't have a label on it. And I guess it makes sense that I would use it consistently with circles. You can multiply it out [*multiply 25 * 3.14]*, but I tend to leave pi as pi. I don't know.
>
> **Bob:** I think it's because I forgot [the units for the circle problem]. Either that or the pi threw me off and then I forgot. Pi doesn't have a unit. I think I forgot because of the unitless pi.

These students say that their lack of unit with the circle and cylinder task was a result of the presence of the symbol π. Our data does not fully explain what is happening with the circle tasks, but the findings do suggest that π being related to unit difficulties warrants further investigation.

Symbolic forms. Sherin's (2001) theory of symbolic forms may explain some of the results of this study, most notably the π and incorrect unit findings. Developed to explain how students understand and construct equations in physics, the symbolic forms framework hypothesizes that students have conceptual schema with which they associate certain symbol patterns in equations. Sherin (2001) developed a list of these symbolic forms, noting that it is not comprehensive and hypothesizing mathematics-specific symbolic forms also exist for mathematics equations. Izsák (2000) explored the existence of a few symbolic forms in students' modeling of physical situations with algebra. It is important to note that in Sherin and Izsák's work, the symbolic form is connected to understandings that students have about the situation they are modeling. That is, a symbolic form and symbol pattern are not something that arise from rote memorization or how something is "supposed to" look: rather, the form and symbol pattern have meaning to the student. For instance, students who understand upwards acceleration and acceleration due to gravity as competing forces might write an equation with a □ - □ symbol pattern because that specific pattern represents a "competing terms" form. Neither Sherin's nor Izsák's accounts of forms include forms regarding units, but our data lead us to believe that (a) such forms exist and (b) they may explain students' unit use. We might say that students understand an area or volume calculation to be measuring something, and because it is a measurement, it has a unit. We propose that there may be a *measurement* symbolic form and an associated symbol pattern as shown in Figure 1. The larger box indicates the magnitude and the smaller box indicates a unit.

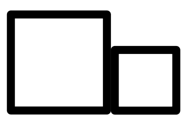


Figure 1. Symbol pattern for "measurement" symbolic form.

If this form exists, it might explain some students' tendency to write a unit with a measurement calculation, even if it is not the correct unit. This form might also explain some student's tendency to include π, but not a unit, for circle and cylinder computations and units on other problems. If "measurement" cues the symbol pattern in Figure 1, π may fill the second box, and students may think the symbol pattern is satisfied. They may then proceed to a different task (say, the volume of the rectangular prism), activate the same measurement schema, and include a unit to satisfy the symbol pattern. Some students might have a form like the one in Figure 2, in which they think of a measurement as represented by a magnitude, the general units of measure (e.g., yards or meters), and the dimension of that unit (e.g., square yards or cubic meters).

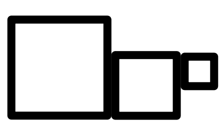


Figure 2. Nuanced symbol pattern for "measurement" symbolic form.

Though further research is needed regarding the existence of these forms, their existence would explain students' unit use and the sometimes contradictory behavior of including units on some problems but not others, or including an incorrect unit. From an instructional viewpoint, this would suggest good news that students do indeed connect area and volume with the measurement concept. We discuss further instructional implications next.

Summary

We found that units of spatial measure cause difficulties for differential calculus students. In particular, only 26.6% of students gave correct units for all of the tasks. One common incorrect unit was a length unit for an area or volume computation. This finding is similar to that of Lehrer (2003), who noted that elementary school students often misappropriate units of length for other spatial measure. We found that students who gave correct units could connect their unit choice to arrays, dimensionality, and rules of exponents. In contrast, students who struggled with units did not seem to have this sort of knowledge. This has implications for teaching area and volume using array models from early grades on, and connecting units to rules of exponents after students have had algebra.

A second important feature of students' responses was not giving units for the circle and cylinder computations. More research is needed to investigate how students think about the symbol π; in particular, we are interested if there might be some connection between the lack of units here and students being instructed in geometry to 'leave your answer in terms of/in units of π' or students being accustomed to things like 3π/2 in trigonometry, where there is a π but no other unit. Finally, further research is needed to investigate our hypothesis about the existence of symbolic forms for students' unit understanding.

References

Battista, M.T., & Clements, D.H. (1998). Students' understandings of three-dimensional cube arrays: Findings from a research and curriculum development project. n R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 227-248). Mahwah, NJ: Erlbaum.

Curry, M., & Outhred, L. (2005). Conceptual understanding of spatial measurement. In P. Clarkson et al. (Eds.), Building Connections: Theory, Research and Practice (Proceedings of the 27th MERGA conference), Sydney: MERGA

Engelke, N. (2008). Developing the solution process for related rate problems using computer simulations. Contributed report, Eleventh Annual SIGMAA on RUME Conference, San Diego, CA, Feb. 29, 2008.

Fuys, D., Geddes, D., & Tischler, R. (1988). The van Hiele model of thinking in geometry among adolescents. *Journal for Research in Mathematics Education Monograph* 3. Reston: NCTM 1988.

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory*. Chicago: Aldine.

Izsák, A. (2005). You have to count the squares: Applying knowledge in pieces to learning rectangular area. *Journal of Learning Sciences, 14,* 361-403.

Izsák, A. (2000). Inscribing the winch: Mechanisms by which students develop knowledge structures for representing the physical world with algebra. *The Journal of the Learning Sciences*, *9*(1), 31-74.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G. Martin, & D. E. Schifter (Eds.), A Research Companion to Principles and Standards for School Mathematics. Reston, VA: National Council of Teachers of Mathematics.

Lehrer, R., Jenkins, M., & Osana, H. (1998). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space (pp. 137-167).* Mahwah, NJ: Erlbaum.

Martin, T. (2000). Calculus students' ability to solve related rates problems. *Mathematics Education Research Journal*, 12(2) p. 74-91.
Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics, 14,* 1-18.
Redish, E.F. (1997). Student difficulties with math in the context of physics. Available online at: http://www.physics.umd.edu/perg/papers/redish/talks/math
Rowland, D.R. (2006). Student difficulties with units in differential equations in modelling contexts. *International Journal of Mathematics, Education, Science, and Technology. 37*(5), p. 553-558.
Rowland, D.R. & Jovanoski, Z. (2004). Student interpretation of the terms in first-order ordinary differential equations in modelling contexts. *International Journal of Mathematics, Education, Science, and Technology. 35*(4), p. 503-516.
Saitta, E. K. H., Gittings, M.J., & Geiger, C. (2011). Learning dimensional analysis through collaboratively working with manipulatives. *J. Chem. Educ. 88*(7), pp. 910 – 915.
Sherin, B. (2001). How students understand physics equations. *Cognition and instruction*, *19*(4), 479-541.

GENERALIZING CALCULUS IDEAS FROM TWO DIMENSIONS TO THREE: HOW MULTIVARIABLE CALCULUS STUDENTS THINK ABOUT DOMAIN AND RANGE

Allison Dorko & Eric Weber
Oregon State University

We analyzed multivariable calculus students' meanings for domain and range and their generalization of that meaning as they reasoned about domain and range of multivariable functions. We found that students' thinking about domain and range fell into three broad categories: input/output, independent/dependent variables, and/or as attached to specific variables. We used Ellis' (2007) actor-oriented generalizations framework to characterize how students generalized their meanings for domain and range from single-variable to multivariable functions. This framework focuses on the process of generalization – what students see as similar between ideas in multiple contexts. We found that students generalized their meanings for domain and range by relating objects, extending their meanings, using general principles and rules, and using/modifying previous ideas. Our results about how students understand and generalize the concepts of domain and range imply that the domain and range of multivariable functions is a topic instructors should explicitly address.

Key words: Calculus, function, generalization

Introduction

This paper focuses on (a) how multivariable calculus students think about domain and range in two and three dimensions and (b) how they generalize their meaning of domain and range from single to multivariable functions. We have two foci because how students generalize their ideas cannot be studied without first identifying what those ideas are. While it is clear to experts that multivariable calculus topics are natural extensions of single-variable calculus topics, how students come to see the relationship between ideas like function and rate of change in single and multivariable contexts is not well understood. Though some recent advances have been made with regard to student thinking about these ideas, these studies are only preliminary (Kabael, 2011; Martinez-Planell & Trigueros, 2013; Trigueros & Martinez-Planell, 2010; Yerushalmy, 1997). Additionally, while there is a large body of knowledge about how students understand various single-variable calculus concepts, far fewer studies exist regarding students' understanding of topics in multivariable calculus. For instance, there is a wide body of knowledge about students' understanding of derivatives of single-variable functions (Asiala, 1997; Orton, 1983; Zandieh, 2000), but not much about students' understanding of derivatives of multivariable functions. This scarcity creates two issues: one, we do not know how students in multivariable calculus think about the concepts presented to them and two, we do not understand how they develop those understandings through the process of generalization.

Gaining insight into these two issues is crucial to many STEM fields, as most mathematics used in the real world involves functions of many variables. For instance, in thermodynamics, energy is a function of pairs of pressure, temperature, volume, and entropy; in engineering, density may be a function of x, y, and z. If the mathematics STEM students are to use involves functions of many variables, it makes sense to study how students understand these functions so that instructors can use that knowledge to address specific difficulties and misconceptions. It is likely that students' understanding of single-variable functions plays a role in their understanding of multivariable functions. Thus this study aimed at not only describing one particular aspect of students' multivariable function understanding, but how that thinking relates to their prior knowledge: in short, what they *see as similar* between domain and range of single and multivariable functions. More broadly, knowing how students generalize in mathematics is useful for instruction in that we can better build on students' prior knowledge and exploit the connections they naturally see between mathematical ideas.

We use domain and range as a 'case study' of how students generalize the meaning of a concept learned with single-variable functions to its meaning for multivariable functions. While domain and range appear in initial instruction about functions, they receive little to no attention in multivariable calculus. For instance, McCallum et al. (2009) do not discuss the domain and range of a function at all. Rowgawski (2008) and Thomas (2010) define and give a few examples of the domains and ranges of multivariable functions. None of these standard texts, however, talk about domain and range in terms of inputs and outputs or independent and dependent quantities, as is commonly done in algebra. Thus most of our subjects had not thought about domain and range in three dimensions, and we were able to observe their initial fits and starts with the ideas and observe detailed and sudden generalizations. This paper centers on the following three organizing themes:

1. What meanings do multivariable calculus students have for domain and range in two dimensions?
2. What meanings do multivariable calculus students have for domain and range in three dimensions?
3. How do multivariable calculus students generalize the concept of domain and range from two dimensions to three dimensions?

Background Literature

There are few articles that discuss students' understanding of domain and range. We searched for articles about students' understanding of domain and range, and when that yielded nothing, we switched to associated terms like 'function machines,' 'input and output,' and 'students' notion of variable'. We searched for 'independence and dependence' in both function literature and statistics education literature. None of these searches resulted in articles that explicitly discuss domain and range, though there are some findings in the function literature related to students' understanding of functions that are relevant to the present study. For instance, one way to define domain and range is the set of inputs and outputs of the function, respectively. According to Oehrtman, Carlson, and Thompson (2008), thinking about a function in terms of an input and corresponding output is the beginning of a robust function conception. Monk (1994) found that most calculus students have developed this pointwise view of function but fewer develop an across-time view of function, in which students' conception of function progress to thinking about the function for infinitely many values and understanding how the a change in one variable affects the other(s). That is, a robust function conception involves not only the ability to pair an input with an output, but an understanding of the relationship between quantities. Confrey and Smith (1995) say the beginning of this understanding occurs as students form connections between values in a function's domain and range. However, as function is introduced in algebra and/or precalculus, the functions instructors ask students to reason about are single-variable functions. How students build an understanding of multivariable functions is not known. Our investigation of students' meanings for domain and rage contributes to the function literature by documenting how students think about domain and range of single- and multivariable functions, and how they generalize the ideas of domain and range.

Generalization

We chose to study this sense making in terms of generalization because the ideas in multivariable calculus are connected to those in single-variable calculus (and, in the case of domain and range, to ideas from algebra), and it is widely believed that students use their prior knowledge in making sense of new topics. More specifically, the ideas in multivariable and single variable calculus are similar and students are likely to pick up on similarities such as terms (e.g., function, domain, range, variable) and symbols (e.g. notations like $f(x)$ and $f(x,y)$; integral symbols). Studying the extension from single to multivariable calculus allows us to see the nature of the connections students make and how they make them. Though there have been many studies about generalization in algebra (e.g. Amit & Klass-Tsirulnikov, 2005; Carpenter & Franke, 2001; Cooper & Warren, 2008; Ellis, 2007), these studies are largely about generalizing patterns, and there are fewer studies of generalization of undergraduate mathematics topics, or studies of the generalization of meaning. As generalization is a critical component of mathematical thinking (Amit & Klass-Tsirulnikov, 2005; Lannin, 2005; Mason, 1996; Peirce, 1902; Sriraman, 2003; Vygotsky, 1986), it is important to extend knowledge of how students generalize in higher mathematics, and in particular how they generalize conceptual meanings.

Theoretical Framework

We studied generalization from an actor-oriented perspective. The actor-oriented perspective attends to what students see as similar in mathematical situations. This is in contrast to an observer-oriented perspective in which students' ideas are compared to what an expert would see as similar across situations. Such perspectives often find that students cannot or do not generalize ideas from one setting to another, and focus on the product – the final general rule or principle – as opposed to the generalization process itself. The actor-oriented perspective allows us to privilege students' perceptions of similarity, and thus their generalization process, even if their perceptions are not necessarily consonant with what an expert would see as similar. We follow Ellis (2007) and Lobato (2003) in thinking about generalization as "the influence of a learner's prior activities on his or her activity in novel situations" (Ellis, 2007, p. 225). This was a useful lens for looking at how students viewed domain and range, a topic they had experienced prior with single-variable functions, in the novel situation of multivariable functions.

Our corresponding analytic framework is Ellis' (2007) generalizations taxonomy. The taxonomy distinguishes between *generalizing actions*, or "learners' mental acts as inferred through the person's activity and talk" (Ellis, 2007, p. 233) and *reflection generalizations*, which are students' public statements about a property or pattern common to two situations. Generalizing actions include relating, searching, and extending (Figure 1). Reflection generalizations include identifications and statements, definitions, and influence (Figure 2). We used this framework to analyze how students generalized their meanings for domain and range.

<table>
<tr><th colspan="3">GENERALIZING ACTIONS</th></tr>
<tr><td rowspan="4">Type I: Relating</td><td rowspan="2">1. Relating situations: The formation of an association between two or more problems or situations.</td><td>Connecting Back: The formation of a connection between a current situation and a previously-encountered situation.</td></tr>
<tr><td>Creating New: The invention of a new situation viewed as similar to an existing situation.</td></tr>
<tr><td rowspan="2">2. Relating objects: The formation of an association between two or more present objects.</td><td>Property: The association of objects by focusing on a property similar to both.</td></tr>
<tr><td>Form: The association of objects by focusing on their similar form.</td></tr>
</table>

Type II: Searching	1. *Searching for the Same Relationship:* The performance of a repeated action in order to detect a stable relationship between two or more objects.
	2. *Searching for the Same Procedure*: The repeated performance of a procedure in order to test whether it remains valid for all cases.
	3. *Searching for the Same Pattern:* The repeated action to check whether a detected pattern remains stable across all cases.
	4. *Searching for the Same Solution or Result:* The performance of a repeated action in order to determine if the outcome of the action is identical every time.
Type III: Extending	1. *Expanding the range of Applicability*: The application of a phenomenon to a larger range of cases than that from which it originated.
	2. *Removing Particulars*: The removal of some contextual details in order to develop a global case.
	3. *Operating*: The act of operating upon an object in order to generate new cases.
	4. *Continuing*: The act of repeating an existing pattern in order to generate new cases.

Figure 1. Generalizing actions for domain and range. Adapted from Ellis (2007).

REFLECTION GENERALIZATIONS		
Type IV: Identification or Statement	1. *Continuing Phenomenon:* The identification of a dynamic property extending beyond a specific instance.	
	2. *Sameness:* Statement of commonality or similarity.	*Common Property*: The identification of the property common to objects or situations.
		Objects or Representations: The identification of objects as similar or identical.
		Situations: The identification of situations as similar or identical.
	3. *General Principle:* A statement of a general phenomenon.	*Rule:* The description of a general formula or fact.
		Pattern: The identification of a general pattern.
		Strategy or Procedure: The description of a method extending beyond a specific case.
		Global Rule: The statement of the meaning of an object or idea.
Type V: Definition	1. *Class of Objects:* The definition of a class of objects all satisfying a given relationship, pattern, or other phenomenon.	
Type VI: Influenc	1. *Prior Idea or Strategy:* The implementation of a previously-developed generalization.	
	2. *Modified Idea or Strategy:* The adaptation of an existing generalization to apply to a new problem or situation.	

Figure 2. Ellis' (2007) reflection generalizations

Data Collection Methods

We interviewed 20 students enrolled in multivariable calculus at a mid-size university in the northwestern U.S. The students were volunteers selected from all the multivariable calculus students enrolled during that term, and were compensated for their participation. The course topics included vectors, vector functions, curves in two and three dimensions, surfaces, partial derivatives, gradients, directional derivatives, and multiple integrals in different coordinate systems. Each student participated in a semi-structured interview that lasted about an hour. We recorded audio and written work from each of the interviews using a LiveScribe Echo Pen, which provides a recording consisting of synced audio and written work. These recordings also allowed us to create dynamic playbacks of the interviews during data analysis. The tasks and rationale for their inclusion are shown in Table 1.

Table 1. Interview tasks and rationale

Task	Rationale
1. What does domain mean? What does range mean?	The purpose of this question was to elicit how students thought about domain and range, and what they associated with the terms, when they were not tied to a specific problem or function.
2. What are the domain and range of $f(x)=4+1/(x-3)$?	This question was included to gain insight into how students operationalized their definitions for domain and range as they worked with a single-variable function.
3. What are the domain and range of $f(x,y)=x^2+y^2$?	This question was included to gain insight into how students thought about domain and range for a multivariable function. We used this task as one way to investigate how students generalized their meanings for domain and range.
4. What are the domain and range of $x^2+y^2+z^2=9$?	This question was included to gain insight into how students thought about domain and range for a multivariable function, and how they thought of domain and range for a function written in a different form than f(x,y).

We had two research foci and thus preformed two separate data analyses. We used a constant comparative analysis (Corbin, 2008) to answer our first two questions, (what are students' meanings for domain and range of single- and multivariable functions?) then did a second analysis using Ellis' (2007) generalizations framework to answer the second (how do students generalize those meanings?). We present the analysis methods, findings, and discussion of students' meanings for domain and range first, then we turn to the analysis, results, and discussion about students' generalizations.

Data Analysis I: Students' Meanings for Domain and Range

We used a constant comparative analysis (Corbin, 2008) to identify what meanings students held for domain and range. Researcher 1, who had done all but two of the interviews, randomly selected half of the interview transcripts and highlighted phrases relating to how students thought about domain and range. Students used words like input, output, result, function as a whole, independent variable, dependent variable, domain goes with x (or x and y), range goes with y (or z), domain goes with the horizontal axis (or plane), range goes with the vertical axis, codomain, and so on. Researcher 1 then read the other half of the transcripts, marking the same words and looking for any other words or phrases students used in thinking about and explaining domain and range. Researcher 1 then looked for themes in this collection of students' phrases, and found that they fit the following categories: (a) Domain and range are associated with specific variable symbols in an equation, (b) Domain and range are inputs and outputs, and (c) Domain and range relate to independent and dependent variables. Researcher 1 created coding criteria for each of these categories for both single and multivariable functions, and both researchers coded all of the data independently. The two researchers compared their results, discussed any differences, and agreed upon the set of codes shown in Table 2. They then used the data within each category to form descriptions of the *meanings* students held for domain and range. In the next section, we give examples of data for each category and describe students' meanings.

Table 2. Codes and criteria for meanings of domain and range

	Criteria	
Code	***Single-Variable***	***Multivariable***
Domain is x, Range is y	Student says that domain is the x values and range is the y values without reference to the notion of function. That is, the student does not mention input, output, independent variable, or dependent variable.	Student is answering a question about $f(x,y)$ and gives a domain for x and a range for y. The student may talk about the $f(x,y)$ or the z value, but still identifies domain as corresponding to x and range as corresponding to y.
Input / Output	Student talks about domain as an input, a value that goes into a function, or a value that "satisfies" the function. Student talks about range as an output value, a 'return value', or the 'result value.' There is a clear reference to the notion of function.	Student is answering a question about $f(x,y)$. Student talks about domain as inputs and identifies that there are multiple inputs because it is a function of more than one variable. Student talks about range as the output, the result of the function, the 'function value,' or the function 'as a whole'. There is a clear reference to the notion of function.
Independence / Dependence	Student identifies that domain corresponds to the independent variable and range corresponds to the dependent variable.	Student is answering a question about $f(x,y)$. Student identifies that domain corresponds to the independent variables and range corresponds to the dependent variable. The student may use the phrase 'determined by' rather than the terms independent / dependent (e.g., "z determined by x and y")

Results & Discussion I: Students' Meanings for Domain and Range of Single and Multivariable Functions

The three broad categories in Table 1 correspond to students' meanings for domain and range. Below, we consider each of these meanings in detail.

Domain is x, range is y

One meaning that students had for domain and range was that domain meant the possible values for x and range means the possible values for y. This meaning was based on the presence of symbols in the equation rather than a notion of function. That is, probing questions about why domain was x or range was y did not yield any underlying explanations of x and y as inputs, outputs, or independent/dependent variable.

The strongest evidence that some students think of domain and range as related to specific symbols is that many students said that the domain was x and the range was y for $f(x,y)$. For instance, Adam and Gabe both defined domain as the possible x values and range as the possible y values for a single-variable function. For $f(x,y) = x^2 + y^2$, they said

Adam: It's a helix, or spinny spring looking thing. Domain and range, so the domain of this would be all real numbers for x values, so x can equal any number, and it changes what z equals, but even negative numbers squared equal positive z. And the range is all real numbers because there is no value of y for which the graph is undefined.

Gabe: So the domain [of $f(x,y) = x^2 + y^2$,] is all real numbers because it's a square so there's no restrictions. And it's the same thing with y, it's the same as the x^2.

Int.: What would it mean if I said 4 is in the domain?

Gabe: You'd just plug it in.

Int.: But do I have to say it for *x* and *y*? If I just say 4 is in my domain and I haven't specified if it's *x* or *y*?

Gabe: I look at the domain as just being *x* values.

Int.: So if I said 4, it would mean that $x = 4$ is in the domain?

Gabe: Yeah.

Int.: What if I made that same statement about the range, if I said 4 was in the range? What would I be looking at?

Gabe: The *y* value.

Adam talked about changes in *x* causing changes in *z*, indicating he understood there was a relationship between *x* and *z*. However, he said that the range was *y*. Thus 'range' seemed to be attached to a specific symbol, rather than the idea of the dependence of one value on another, as he had mentioned earlier. Gabe associated domain with only *x* values, and range with only *y* values for both *f*(x) and *f*(x,y). In summary, the meaning of domain and range for these students was that domain corresponds to *x* and range corresponds to *y*, whether the function was a single variable function or a multivariable one.

Input/Output

One way to think about function is as a machine that takes inputs and returns outputs. Many students thought about domain and range as related to this notion of function. To these students, domain meant the possible inputs to a function and range meant the possible outputs. For single-variable functions, students identified a singular output. For multivariable functions, students explained domain as corresponding to multiple inputs, as Jim did by identifying an *x* input and a *y* input. The input-output meaning often included a link between the inputs and outputs, such that each choice of input produced a particular output. For instance, Deb talked about 'return values for each *x* in the domain'.

Deb: In terms of $f(x) = y$, domain would be all the value that go into the function. The domain will be all of the values for *x* that return a unique, I think, value for *y*. The range would be all the return values for each *x* in the domain.

Jim: Domain is your input values, otherwise known as your *x* values. It could also represent your independent values. The range is your output, your dependent variables, *y* values.
[Q3] There would be two different domains. You have your *x* input and your *y* input. Your *x* domain and your *y* domain give you a range of a different variable. It's the range of *z* or *f*(x,y).

Note that Jim talked about both inputs/outputs and independence and dependence, so his answer was coded as belonging to both categories. It was fairly common for students to understand domain and range in terms of both input and output and independence and dependence.

Independence / Dependence

A function may be thought of as a relationship in which the value of one variable depends on the value of another variable. For students who thought about function this way, domain meant the possible values of the independent variable and range meant the possible values of the dependent variable. Kathy gave a good example of this with equations, and Leah talked about independence and dependence graphically by thinking about a "*y* plane" as determined by *x* values. Both Leah and Phillip identified that a multivariable function has multiple independent variables.

Kathy: Domain is the range of *x* values that a function can have. And I guess *x* is just the independent variable. If the function were *f*(y), the domain would be *y*. Range is the values that a function has for the given domain. Usually it's f(x) = y. Then y has the range.

Leah: Domain is the range of values the dependent variable can take. No, it's the independent. It's the y plane determined by the x value, or the z determined by the x and y.

Phillip: [Q3] It's a function of two variables. X and y are both independent variables, rather than the dependent variable. You could say the domain is the independent variable and range is the dependent variable.

In summary, the meanings students held for domain and range included 'domain is x and range is y'; domain as input and range as output; and domain and range as related to independent and dependent variables. In the next section, we describe how we analyzed students' *generalizations* of these meanings from their meaning in f(x) to their meaning in f(x,y).

Data Analysis II: Coding Students' Generalizations

Our second analysis was to determine how students generalized their meanings for domain and range as they moved from working with f(x) to thinking about f(x,y). We based this analysis on Ellis' (2007) generalizations framework. The framework distinguishes between *generalizing actions*, which are "students' activity as they generalize" (Ellis, 2007, p. 198), and *reflection generalizations*, which are "final statements of generalization (verbal or written) or the use of a result of a prior generalization" (Ellis, 2007, p. 198). In the next subsections, we explain how we used this framework to code our own data.

Generalizing Action: Relating

Relating is a generalizing action in which "students form an association between two or more problems, situations, ideas, or mathematical objects. They relate by recalling a prior situation, inventing a new one, or focusing on similar properties or forms of mathematical objects" (Ellis, 2007, p. 198). We only found two instances of *relating situations*. One student who defined domain and range as relating to independent and dependent variables *connected back* to a physics lab in which an experiment had had such variables. A different student, who defined domain and range in terms of inputs and outputs, engaged in *creating new* by describing temperature in California as a function of temperature in Oregon, and explained that the temperature in Oregon would be the input.

Relating objects was far more common. We found that students related both equations and graphs or coordinate axes. For instance, both Leah and Mimi related the coordinate axes of R^2 to the coordinate axes of R^3:

Leah: Range is the y plane determined by the x value, or z determined by x and y. [*Relating objects: property*]

Phillip: Lets call z the dependent variable here and move the x and y to the other side. Now the domain is x and y. [*Relating objects: property*]

Mimi: You can't have negative z but I don't know if that's the domain or the range. I'm going to say it's the range, and treat the z axis like the y axis of the function. [*Relating objects: form*]

Leah and Phillip related the coordinate axes based on the property of independence and dependence, which Leah called 'determined by.' Mimi did not use a mathematical property to relate the axes, but instead seemed to see as similar the vertical position of the y axis in f(x) and the z axis in f(x,y).

One clear instance of relating objects by their form was the category of students who said that domain was x and range was y for both f(x) and f(x,y). In these cases, the presence of x and y in an equation seemed to trigger students to say that domain was the possible x values and range was the possible y values. Ian and Gabe's descriptions of domain and range are good examples:

Ian: [Q1] [Domain] is whatever the x value can be. The values the x component can be composed of. [Range] would pretty much be the same thing except for the y component.

[Q3] so whatever x is, it would be whatever values z is because that would be the radius [writes 'domain: $-z < x < z$']. And the y is the same [writes 'range: $- y < z < y$'].

[*Relating objects: form]*

Gabe: So the domain [of f(x,y) = $x^2 + y^2$,] is all real numbers because it's a square so there's no restrictions. And it's the same thing with y, it's the same as the x^2.

Int.: What would it mean if I said 4 is in the domain?

Gabe: You'd just plug it in.

Int.: But do I have to say it for x and y? If I just say 4 is in my domain and I haven't specified if it's x or y?

Gabe: I look at the domain as just being x values.

Int.: So if I said 4, it would mean that $x = 4$ is in the domain?

Gabe: Yeah.

Int.: What if I made that same statement about the range, if I said 4 was in the range? What would I be looking at?

Gabe: The y value.

For Ian and Gabe, domain meant x and range meant y. Thus what they saw as similar in f(x) and f(x,y) was that both had an x and a y. They generalized their meaning for domain and range based on the presence of the variables in the equation. This was true of all students in the 'Domain is x, Range is y' category: students who thought domain was x and range was y in both single-and multivariable functions seemed to have made that generalization based on the presence of the variables in the equations rather than based on a conceptual meaning for domain and range.

Generalizing Action: Extending

Ellis (2007) defines *extending* as a generalizing action that "involves the expansion of a pattern, relationship, or rule into a more general structure. Students who extend widen their reasoning beyond the problem, situation, or case in which it originated" (Ellis, 2007, p.198). Our students extended the range of applicability and removed particulars. The following excerpts are representative of the ways in which students engaged in extending.

Jim: [Q1] Domain is your input values, otherwise known as your x values. It could also represent your independent values. The range is your output, your dependent values, your y values.

[Q3] There would be two different domains because there are two different inputs. I guess the range could be any number just dependent on the domain, like you could put anything into the domain and you would get a range number out. Your x domain and your y domain give you a range of a different variable. So it would be, the range would be of f(x,y).

[*Extending: removing particulars*]

Bailey: I think in 2 dimensions, whatever your domain is, you put that in and that's what your output is. I suppose that's the same in 3D as well: the array of possible values I can get out of the function.

[*Extending: removing particulars*]

Deb: [Q1] The domain is all the values for x that return a unique value for y. The range would be all of the return values. In 3D, the domain is all values for x and y and the range is all values for z.

[Q4] I am going to use a graph because I know it's a sphere. So the domain would be all the values between… it's like R but it's kind of limited between 3 and -3 on each part. So -3 to 3 for x, y, z. Those are domains. The range, it won't be 3 any more because we have… I am not sure about the range. What are the return values. I'll write it as $z = \sqrt{9 - x^2 - y^2}$. Now the range would be, that is R. [*Extending*: *expanding the range of applicability*]

What Jim saw as similar between the domain of f(x) and f(x,y) was that in each case, domain meant input. He thus extended his idea of domain-as-input to domain-as-inputs, and likewise extended the idea of 'getting a range number out' to f(x,y) representing that number just as f(x) did. We coded this as *extending: removing particulars* because Jim removed the contextual details of the problem (that is, the function $f(x,y) = x^2 + y^2$) in order to develop a global case: domain is the input(s) and range is the output. He put the actual equation while foregrounding the meaning of domain and range. Likewise, Bailey extended the idea of range being "the array of possible values I can get out of the function" to decide that range was "the same in 3D." In stating this, she removed the particulars of the specific equation as Jim had. Deb also removed particulars, extending the idea of range as a "return value" when she worked with the equation for the sphere. Deb's meaning for range in 2D had been a return value or a z value. However, the equation for the sphere was written differently than the other equations. Deb extended by asking herself what the return value was, then solved the equation for z so she could apply her meaning for range. In doing so, she extended the range of applicability because she applied a meaning to something different from which it had originated.

Reflection generalization: Identification or statement: General principle

Ellis (2007) defines a *general principle* as "a statement of a general phenomenon" (Ellis, 2007, p. 200). General principles come under the categories of 'identification or statement' in which students make their generalizations public by explicitly writing or stating them.. Our students frequently stated global rules as they tried to think about the meaning of the domain and range of f(x,y). That is, one way in which they made meaning of the concepts "domain of f(x,y)" and "range of f(x,y)" was to state their meaning of the concepts "domain of f(x)" and "range of f(x)", linking the meaning in each context to form a description of the general phenomenon. For example,

Mimi: Like you've got x, you've got y, and z is kind of like the function value. It equals f(x,y) kind of like $y = f(x)$. It's the dependent variable, not the independent.

Philip: The range… is the result of the function, so I guess that would be z. The range is … the dependent variable. X and y are both independent variables. You could give a better definition than in question 1 and say domain is the independent variable and range is the dependent variable.

Mimi and Phillip used two ideas in their meaning of range: that of the "function value" or "result of the function" and that of dependency. The function value meaning allowed Mimi to see $z = f(x,y)$ as analogous to $y = f(x)$. Likewise, Phillip saw z as the "result" of the function of x and y. He stated a global rule that domain corresponds to the independent variables and range corresponds to the dependent variable. In talking about the function's value or result and independence/dependence, the students were stating the *meaning* of domain and range.

Phillip's statement is a good example of the relationship between generalizing actions and reflection generalizations. Ellis (2007) notes that reflection generalizations often come on the heels of generalizing actions. Phillip *extended* his idea about "the result of a function" from the single-variable to the multivariable case, and this extension was immediately followed by a synthesizing comment about the meaning of domain and range in general.

Reflection Generalization: Influence

There are two reflection generalizations classified as Influence. The first is *prior idea or strategy*, in which a student implements a previously developed generalization. The second is a *modified idea or strategy*, in which a student adapts an existing generalization to apply to a new problem or situation. Quincy and Neil's statements illustrate the difference well:

Quincy: [Q1] Range is how far the function spans. Range is the set of numbers the function can have.

[Q4] I think the range is 9 for this one… because that's the value on the other side of the equal sign. So it can't really range to any other values.

Neil: [Q1] Domain is the span that the x value can take on. Range is the span that the y value can take on.

[Q3] In this instance the range is z, the output value. So I would say the variables applied to the function doesn't necessarily correspond to domain as x, range as y. So if I looked back to my definitions in question one, I could define domain and range in 3D space with domain as the span of values that can occur on the horizontal plane and I would define range to be the span of values that are dependent on the domain and span the vertical plane.

Quincy directly applied his generalization that "range is the set of numbers the function can have" to the equation for the sphere, noting that the only number the x, y, and z could add to was 9. Thus the "set" of numbers that function had consisted of one element (namely, 9). In contrast to Quincy, who implemented an existing generalization, Neil modified his existing generalization that domain was x and range was y. Since that generalization did not seem to apply to $f(x,y) = x^2 + y^2$, he adapted his idea such that to domain was the horizontal plane and range was a dependent quantity, illustrated graphically as the vertical plane.

Results & Discussion II: How Students Generalize Their Meanings for Domain and Range

We found that students generalize their meanings for domain and range by relating situations, relating objects, and extending their meanings beyond the cases in which they had originated. However, our students did not engage in all of the generalizing actions or reflection generalization that Ellis (2007) identifies. We think that this is likely an artifact of how the data were collected: Ellis' data come from a problem-based teaching experiment focused on deriving linear relationships, while our data comes from a single interview.

Table 3. Generalizing actions for domain and range.

Ellis (2007) framework			**Example in domain/range data**
Type I: Relating	1. Relating situations: The formation of an association between two or more problems or situations.	Connecting Back: The formation of a connection between a current situation and a previously-encountered situation.	Domain is your input values. It could also represent your independent values. I am trying to think like in terms of my physics lab where there are independent and dependent variables and you plug in the numbers that you use.
		Creating New: The invention of a new situation viewed as similar to an existing situation.	Say you need to calculate temperature and you have the temperature relative to California and you have some conversion, so the input values are the temperatures in Oregon and the output values are the temperature in California.

	2. Relating objects: The formation of an association between two or more present objects.	Property: The association of objects by focusing on a property similar to both.	Lets call *z* the dependent variable here and move the *x* and *y* to the other side. Now the domain is *x* and *y*.
		Form: The association of objects by focusing on their similar form.	You can't have negative *z* but I don't know if that's the domain or the range. I'm going to say it's the range, and treat the *z* axis like the *y* axis of the function.
Type III: Extending	1. Expanding the range of Applicability: The application of a phenomenon to a larger range of cases than that from which it originated.		Domain is your input values, otherwise known as your *x* values. It could also represent your independent values. The range is your output, your dependent values, your *y* values.
	2. Removing Particulars: The removal of some contextual details in order to develop a global case.		I am a little fuzzy on range in 3D. I think in 2 dimensions, whatever your domain is, you put that in and that's what your output is. I suppose that's the same in 3D as well: the array of possible values I can get out of the function.

Relating objects was a common way that students generalized their meanings of domain and range. When students related equations, some saw the symbols *f*(x) and *f*(x,y) as serving a similar purpose in the equation, namely as the output or the "result" of the function. This allowed them to justify that range, meaning the output or result of a function, would apply to *f*(x,y). Others related coordinate axes, some incorporating an independence/dependence meaning (e.g., Leah's *y* axis determined by the *x* axis and *z* axis determined by the *xy* plane) and others seeming to see as similar the axes' orientation in space (e.g., range applies to whatever axis is vertical and domain to whatever axes are horizontal). A final relation of objects was students' seeing as similar that both *f*(x) equations and *f*(x,y) equations contained the same variables. Students who used this relation often said that the domain of *f*(x,y) was *x* and the range was *y* because that was true for *f*(x,y).

Our students also generalized by extending their meanings of domain and range in the single-variable case to the multivariable case. These extensions often involved expanding the range of applicability, such as extending the ideas of an independent *x* and a dependent *y* to an independent *x* and *y* and a dependent *z* or extending the idea of an input *x* and an output *y* to an input of *x* and *y* and an output *z*. For some students, extending involved removing particulars (like the actual equation) to focus on the meaning of domain and range (e.g., as input and output). When students extend, they place in the background the equations they are reasoning about and foreground the meaning of the concepts.

The reflection generalizations our students stated came in the form of general principles, prior ideas, and modified ideas. Ellis (2007) notes that students' reflection generalizations often mirror their generalizing actions, and it makes sense that our students' extensions (generalizing actions) often resulted in statements of global rules, or statements in which they used or adapted a previous generalization to incorporate the new case of multivariable functions. As with generalizing actions, not all of Ellis' (2007) categories for reflection generalizations were present in our data. The omissions are *continuing phenomena*, *sameness*, and *definition*. The reflection generalization taxonomy for these data are in Table 4.

Table 4. Reflection generalizations for domain and range.

<table>
<tr><th colspan="3">Ellis (2007) framework</th><th>Example in domain/range data</th></tr>
<tr><td rowspan="2">Type IV: Identification or Statement</td><td rowspan="2">3. General Principle: A statement of a general phenomenon.</td><td>Rule: The description of a general formula or fact.</td><td>[Q3a] Domain of this would be all real numbers for x values, so x can equal any number, and it changes what z equals, but even negative numbers squared equal positive z. And the range is all real numbers because there is no value of y for which the graph is undefined.</td></tr>
<tr><td>Global Rule: The statement of the meaning of an object or idea.</td><td>Z is kind of like the function value. It equals $f(x,y)$ kind of like $y = f(x)$. It’s the dependent variable, not the independent.</td></tr>
<tr><td rowspan="2">Type VI: Influence</td><td colspan="2">1. Prior Idea or Strategy: The implementation of a previously-developed generalization.</td><td>[Q1] Range is the set of numbers the function can have. [Q3b] I think the range is 9 for this one… because that's the value on the other side of the equal sign. So it can't range to any other values.</td></tr>
<tr><td colspan="2">2. Modified Idea or Strategy: Z is kind of like the function value. It equals $f(x,y)$ kind of like $y = f(x)$. It’s the dependent variable, not the independent.</td><td>In this instance the range is z, the output value. So I would say the variables applied to the function doesn’t necessarily correspond to domain as x, range as y. So if I looked back to my definitions in question one, I could define domain and range in 3D space with domain as the span of values that can occur on the horizontal plane and I would define range to be the span of values that are dependent on the domain and span the vertical plane.</td></tr>
</table>

That our data contained many of Ellis’ (2007) categories for how students generalize supports the framework as useful for analyzing students’ generalizations.

Implications for Instruction

Devoting Time to Domain and Range

The actor-oriented transfer theoretical framework is useful for exploring generalization because it characterizes what students see as similar without comparing students’ perspectives to those of experts. However, judging whether students’ generalizations are congruent with experts’ ideas becomes useful when thinking about implications for instruction. For instance, some of the ways in which students related objects allowed them to generalize that the domain of $f(x,y) = z$ was x and y and the range was z. Students who formed this generalization commonly used a meaning for domain and range as input and output or independent and dependent variables along with their generalizing action of relating objects. In contrast, students who generalized incorrectly –(relating $f(x) = y$ to $f(x,y) = z$ by concluding that the x and y were present in both equations, and thus played the same role in both) seemed to not have a conceptual meaning for domain and range, but rather a definition that was a link between a word and a symbol (that is, x is domain, y is range). As it seems to be the underlying meaning the first set of students had that allowed them to relate objects in a productive way, instructors might focus on the meaning of $f(x,y)$ as a function with multiple inputs, similar to $f(x)$ (a function with one input). Many of our students stated that the interview was the first time they had thought about the domain and range of multivariable functions. Given some students' incorrect generalizations, it would likely be beneficial for instructors to devote time to talking about the domain and range of multivariable functions.

Complementing with a Focus on Covariation

We also recognize that a strong notion of input and output is not necessarily enough for students to think about function in the ways instructors intend. A generalized notion of input and output has limitations because it relies on the notion that one quantity is dependent on another. In most real world situations, the notion of independence and dependence is contrived because one quantity's value is not actually determined by another quantity's value. While it may be useful to treat one quantity as dependent for ease of calculation of simplification of some physical situation, thinking about functions in terms of covariation is crucial to students' success in calculus (Thompson & Silverman, 2008). In short, thinking covariationally means the student thinks about a function as an invariant relationship between quantities' values not necessarily coupled with a notion of input and output (Thompson, 2011). For example, consider a situation in which a person is moving and there are two quantities: the amount of distance she has traveled and the amount of time elapsed since she began traveling. One would be reasoning covariationally if a) she conceived of both quantities and their individual variation (i.e. time varies, distance varies) and b) she conceived of those quantities varying simultaneously, so that when she thinks about a person's distance traveled, she has an image of the amount of time needed to travel that distance. There is no sense of input or output *required* (though it may be present) within covariational reasoning. Inputs, outputs, independence, and dependence ideas may (i.e. elapsed time causes elapsed distance, or vice versa) arise because of the person's conception of the situation, not because one quantity has been designated as an input and one as an output. It is important to note that covariational reasoning does not preclude an approach involving input and output. Instead, it focuses on a quantitative relationship as the basis for a function from which an input-output metaphor may or may not be drawn. Thus, while this study shows ways in which one might generalize notions of input and output, it is important that multivariable functions not be presented and talked about solely in terms of input and output. While it maybe a useful way to think about domain and range, it does not guarantee that students think about functions as they need to (that is, in terms of covariation) as is useful for calculus.

Suggestions for Further Research

Our tasks included functions of one and two variables. It would be interesting to include functions of more than two variables, such as f(w,x,y,z). A task including this might yield interesting results with students who have the 'variable perspective' (i.e., domain is x and range is y) as they must now think about variables which do not appear in $f(x) = y$. That is, the symbol w does not appear in this equation and thus as students try to explain its place in f(w,x,y,z), they might reveal things about their concepts of domain and range which were not revealed in our tasks.

This study was done with multivariable calculus students, but the concepts of domain and range are used in mathematics outside of calculus. For instance, domain and range are critical in linear transformations. Thus how linear algebra students generalize ideas of domain and range would provide an additional opportunity to study generalization, as well as the meanings for domain and range students have after a higher mathematics course.

Finally, as noted earlier, domain and range were a 'case study' of generalization in higher mathematics. There are many more single- and multivariable calculus ideas in which to explore students' generalizations; of particular interest to us are how students generalize ideas of derivatives and integration.

References

Asiala, M., Cottrill, J., Dubinsky, E., & Schwingendorf, K. E. (1997). The development of students' graphical understanding of the derivative. . (1997). The development of students' graphical understanding of the derivative. *The Journal of Mathematical Behavior, 16*(4), 399-431.

Confrey, J. S., E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education, 26*(1), 66-86.

Corbin, J. S., A. L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. CA: Sage Publications

Ellis, A. B. (2007). A taxonomy for categorizing generalizations: generalizing actions and reflection generalizations. *The Journal of the Learning Sciences, 16*(2), 221-262.

Kabael, T. U. (2011). Generalizing single variable functions to two-variable functions, function machine and APOS. *Educational Sciences: Theory and Practice, 11*(1), 16.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17-20.

Martinez-Planell, R., & Trigueros, M. (2013). Graphs of functions of two variables: results from the design of instruction. *International Journal of Mathematical Education in Science and Technology, 44*, 663-672.

McCallum, W. G., Hughes-Hallet, D., Flath, D., Osgood, B.G., Gleason, A.M., Quinney, D., Frazer Lock, P., Tecosky-Feldman, J., Mumford, D., & Tucker, T.W. (2009). *Multivariable Calculus* (5th Edition ed.). Hoboken, NJ: Wiley.

Monk, G. S. (1994). Students' understanding of function in calculus courses. *The Humanistic Mathematics Network Journal, 27*, 21-27.

Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). *Foundational reasoning abilities that promote coherence in students' understandings of function* (M. P. C. C. Rasmussen Ed.). Washington, DC: Mathematical Association of America.

Orton, A. (1983). Students' understanding of differentiation. *Educational Studies in Mathematics, 14*(3), 235-250.

Rowgawski, J. (2008). *Calculus: Early Transcendentals* New York, NY: W.H. Freeman and Company.

Thomas, G. B. (2010). *Calculus: Early Transcendentals*. Boston, MA: Pearson.

Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education* (pp. 33-57). Laramie, WY: University of Wyoming.

Thompson, P.W. & Silverman, J. (2008). The concept of accumulation in calculus. In M.P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research in teaching in undergraduate mathematics* (pp.43-52). Washington, DC: Mathematical Association of America.

Trigueros, M., & Martinez-Planell, R. (2010). Geometrical representations in the learning of two-variable functions. *Educational Studies in Mathematics, 73*, 3-19.

Yerushalmy, M. (1997). Designing representations: Reasoning about functions of two variables. *Journal for Research in Mathematics Education, 28*(4), 431-466.

Zandieh, M. J. (2000). *A theoretical framework for analyzing student understanding of the concept of derivative*.

Graduate students Teaching Assistants' (GTAs') beliefs, instructional practices, and student success

Jessica Ellis
San Diego State University

In this report I present findings from a large, national study focused on Calculus I instruction. Graduate student Teaching Assistants (GTAs) contribute to Calculus I instruction in two ways: : as the primary teacher and as recitation leaders. As teachers, GTAs are completely in charge of the course just as a lecturer or tenured track/ tenured faculty would be, although they lack the experience, education, or time commitment of their faculty counterparts. In this study, I investigate how GTAs compare to tenure track/tenured faculty, and other full/part time faculty on their (a) beliefs about mathematics; (b) instructional practices; and (c) students' success in Calculus I. Findings from this report point clearly to a need to prepare GTAs adequately for the teaching of calculus but also for further examination of the nature and implications of the differences between GTA and other instructor types' beliefs about teaching and teaching practices.

Keywords: *Graduate student Teaching Assistants (GTAs), Calculus instruction, beliefs, instructional practices, student success*

In this study I investigate the relationship between Graduate Student Teaching Assistants (GTAs) and various aspects of Calculus I instruction. Graduate student Teaching Assistants contribute to Calculus instruction in two ways: as the primary teacher and as recitation leaders. As teachers, GTAs are completely in charge of the course just as a lecturer or tenured track/ tenured faculty would be, although they lack the experience, education, or time commitment of their faculty counterparts. In the College Board of Mathematical Sciences (CBMS) 2010 report, GTAs were found to have taught seven percent of the 234,000 students enrolled in mainstream Calculus I, and 17% of all mainstream Calculus I sections at PhD institutions (Blair, Kirkman, & Maxwell, 2012). Mainstream calculus refers to the calculus course that is designed to prepare students for the study of engineering or the mathematical or physical sciences. In this report, a course was reported to be taught by a GTA only when the GTA was the instructor on record. Thus, these numbers exclude the GTAs who ran discussion or recitation sections.

GTAs can also be viewed as the next generation of mathematics instructors. Thus, in addition to their immediate contribution to the landscape of Calculus I instruction, GTAs contribute significantly to the long-term state of Calculus. The preparation GTAs receive to prepare them for teaching Calculus therefore influences both their immediate teaching practices as well as their long-term pedagogical approach. There has been much discussion about what knowledge and experiences are needed to foster excellent (or even adequate) teachers in mathematics at the K-12 level (Ball, Thames, & Phelps, 2008; Hill, Ball, & Schilling, 2008; Shulman, 1986) and instructors at the undergraduate level (Johnson & Larsen, 2012; Speer, Gutmann, & Murphy, 2005). From these discussions, it is clear that expertise in mathematics alone is not sufficient in the preparation of teachers. Professional development efforts to improve teaching are often aimed at developing teachers' knowledge, beliefs, and instructional practices in order to improve their students' success and to enculturate new teachers into the teaching community (Putnam & Borko, 2000; Sowder, 2007). However, little is known about how GTAs' compare to other instructor types along these dimensions. Accordingly, I have identified the following research question: *How do GTAs compare to tenure track/tenured faculty, and other full/part time faculty on their (a) beliefs about mathematics; (b) instructional practices; and (c) students' success in Calculus I?*

Research Methodology

To answer this question, I draw upon data coming from a large, nationwide study focused on successful calculus programs: Characteristics of Successful Programs in College Calculus (CSPCC). The first phase of the CSPCC study comprised of six surveys: three surveys given to students (one at the beginning of Calculus I, one at the end of Calculus I, and one a year later), two surveys given to instructors (one at the beginning of Calculus I and one at the end of Calculus I), and one survey given to the Calculus course coordinator. The surveys were sent to a stratified random sample of mathematics departments following the selection criteria used by Conference Board of the Mathematical Sciences (CBMS) in their 2005 Study (Lutzer et al, 2007). For the purposes of surveying post-secondary mathematics programs in the United States, the CBMS separates colleges and universities into four types, characterized by the highest mathematics degree that is offered: Associate's degree (hereafter referred to as two-year colleges), Bachelor's degree (referred to as undergraduate colleges), Master's degree (referred to as regional universities), and Doctorate (referred to as national universities). Within each type of institution, we further divided the strata by the number of enrolled full time equivalent undergraduate students, creating from four to eight substrata. Institutions with the largest enrollments were sampled most heavily. In all, we selected 521 colleges and universities: 18% of the two-year colleges, 13% of the undergraduate colleges, 33% of the regional universities, and 61% of the national universities. Of these, 222 participated: 64 two-year colleges (31% of those asked to participate), 59 undergraduate colleges (44%), 26 regional universities (43%), and 73 national universities (61%).

The goals of these surveys were to gain an overview of the various calculus programs nationwide, and to determine which institutions had successful calculus. *Success* was defined by a combination of student variables: persistence in Calculus as marked by stated intention to take Calculus II; affective changes, including enjoyment of math, confidence in mathematical ability, interest to continue studying math; and passing rates. These variables will be used to discuss student success. The instructor surveys address various components of instructors' knowledge, espoused beliefs, and instructional practices. The course coordinator survey addresses programmatic qualities that can be used to situate the individual GTAs within their institutions as well as to gain a topical understanding of the training and support structures available to GTAs, as stated by their course coordinators.

There were 535 instructors who responded to one of the surveys linked to 6306 students, coming from 136 institutions. As shown in Table 1, 30% of the instructors came from a large national university (over 20,000 students) and taught 35% of the students, 30% from a small national university (less than 20,000 students) and taught 30% of the students, 10% from a regional university and taught 6% of the students, 18% from a undergraduate college and taught 22% of the students, and 13% from a two-year college and taught 6% of the students. As shown in Table 2, 46% of the instructors reported being tenure track or tenured and taught 40% of the students, 37% reporting to be "other full or part time faculty" and taught 48% of the students, and 17% report being GTAs who taught 12% of the students. GTAs only taught at national universities, with 67% at large national universities.

Table 1. Number of instructors and students from each institution type.

Institution Type	Instructors	%	Students	%
Two-year colleges	68	12.7	365	5.8
Undergraduate colleges	96	17.9	1381	21.9
Regional universities	54	10.1	377	6.0
Small national universities (<20,000)	156	29.2	1940	30.8
Large national universities (>20,000)	161	30.1	2243	35.6
Total	*535*		*6306*	

Table 2. Number of instructors and students from each instructor type.

Instructor Type	Instructors	%	Students	%
Tenure track/ Tenured faculty	246	46.0	2477	39.3
Other full or part time faculty	197	36.8	3052	48.4
Graduate teaching assistant	92	17.2	777	12.3
Total	*535*		*6306*	

I answer the research question by conducting descriptive analyses to determine differences between instructor types (tenure track/tenured faculty and other full/part time faculty) across a number of variables, addressing knowledge and beliefs about mathematics, instructional practices, and student success.

Results

Beliefs about doing, teaching, and learning mathematics

The first dimension of teaching practice that I compare between tenure/ tenure track faculty, other full and part time faculty, and GTAs is their beliefs about doing, teaching, and learning mathematics. As shown in Table 3, there were significant differences between the types of instructors for three of the reported beliefs about teaching mathematics: “graphing calculators or computers help students understand underlying mathematical ideas (1) or find answers to problems” (6) [$F(2, 516) = 4.193$, $p = .016$], and “all students in beginning calculus are capable of understanding the ideas of calculus” [$F(2, 389) = 3.112$, $p = .046$], and “if I had a choice, I would continue to teach calculus” [$F(2, 385) = 5.969$, $p = .003$]. For all other beliefs about doing teaching, or learning mathematics, there were no significant differences between reported frequencies based on instructor type.

Post hoc comparisons using the Tukey HSD test indicated that the mean response to the prompt “graphing calculators or computers help students understand underlying mathematical ideas (1) or find answers to problems (6)” was significantly different between tenure track/ tenured track faculty (3.62, 1.62) and GTAs (4.17, 1.55), but there were no significant differences between the other types of instructors. The mean response to the prompt “all students in beginning calculus are capable of understanding the ideas of calculus” was significantly different between tenure track/ tenured track faculty (3.63, 1.52) and GTAs (4.15, 1.27), but there were no significant differences between the other types of instructors. Finally, the mean response to the prompt "if I had a choice, I would continue to teach calculus" was significantly different between tenure track/ tenured track faculty (5.19, 1.04, 1.68) and GTAs (4.66, 1.25), and between other full or part time instructors (5.14, 1.09) and GTAs, but not between full or part time instructors and tenure track/ tenured faculty.

These results indicate that GTAs believe that technology serves as a procedural aid more than a conceptual aid when compared to tenure/tenure track faculty, that GTAs view their students as more capable of understanding calculus than tenure/tenure track faculty, and GTAs are slightly less interested in teaching calculus than all other types of instructors. These results also indicate that GTAs report holding similar beliefs about doing, teaching, and learning mathematics for all others beliefs questions.

Table 3. Beliefs about doing, teaching, and learning mathematics by instructor type.

Belief about doing, teaching, or learning mathematics:	**Tenure track/ Tenured faculty**	**Other full or part time faculty**	**Graduate teaching assistant**
From your perspective, in solving Calculus I problems, graphing calculators or computers help students:** *(1=understand underlying mathematical ideas; 6=find answers to problems)*	3.62 (1.62)	3.81 (1.45)	4.17 (1.55)
All students in beginning calculus are capable of understanding the ideas of calculus.** *(1=strongly disagree; 6=strongly agree)*	3.63 (1.52)	3.70 (1.51)	4.15 (1.27)
If I had a choice, I would continue to teach calculus.** *(1=strongly disagree; 6=strongly agree)*	5.19 (1.04)	5.14 (1.09)	4.66 (1.25)

Note.* = p $\leq$ .10, ** = p $\leq$ .05, *** = p $\leq$ .001; Standard deviation in parentheses.

Instructional Practices

As shown in Table 4, there were significant differences between the types of instructors for four of the reported instructional activities: having students work with one another [F(2, 404) = 6.084, p = .002], holding a whole-class discussion [F(2, 403) = 2.495, p = .084], and having students give presentations [F(2, 400) = 3.927, p = .020]. For all other instructional activities, there were no significant differences between reported frequencies based on instructor type.

Post hoc comparisons using the Tukey HSD test indicated that the mean frequency for having students work with one another was significantly different between tenure track/ tenured track faculty (2.80, 1.68) and other full or part time faculty (3.24, 1.73), and between tenure track/ tenured track faculty and GTAs (3.59, 1.86). The mean frequency for holding whole class discussion was significantly different between other full or part time faculty (3.20, 1.76) and GTAs (2.69, 1.26), but there were no significant differences between the other types of instructors. Finally, the mean frequency for having students give presentations was significantly different between tenure track/ tenured faculty (1.46, .96) and GTAs (1.87, 1.25), but there were no significant differences between the other types of instructors.

These results indicate that GTAs report having students work together significantly more frequently than tenure track and tenured faculty, holding whole class discussion significantly less frequently than other full and part time faculty, and have students give presentations significantly more frequently than tenure track and tenured faculty. Taken together, these results indicate that GTAs report different instructional practices than tenure track tenured and other full and part time faculty.

Table 4. Instructional practices by instructor type.

During class, how frequently did you:	**Tenure track/ Tenured faculty**	**Other full or part time faculty**	**Graduate teaching assistant**
(a) show students how to work specific problems?	5.18 (1.14)	5.28 (.99)	5.14 (.97)
(b) have students work with one another? **	2.80 (1.68)	3.24 (1.73)	3.59 (1.86)
(c) hold a whole-class discussion? **	3.13 (1.69)	3.20 (1.71)	2.69 (1.26)
(d) have students give presentations? *	1.46 (.96)	1.68 (1.23)	1.87 (1.25)

Note.* = p ≤ .10, ** = p ≤ .05, *** = p ≤ .001; Standard deviation in parentheses.

Student success

The final dimension that I compare GTAs to other instructor types on is their students' success. In order to measure student success in Calculus I, I used five variables: persistence onto Calculus II, expected pass rate, and three affective measures – change in confidence in mathematical ability, change in enjoyment in doing mathematics, and increased interest in taking mathematics. These measures of success were chosen because many students enter Calculus I pursuing a STEM degree and change their major away from a STEM field because of a decreased interest or enjoyment in mathematics. Research into the reasons students switch out of STEM majors consistently points to the calculus classroom environment as the underlying commonality (Rasmussen & Ellis, 2013; Seymour & Hewitt, 1997; Thompson et al., 2007). As shown in Table 5, there are significant differences in the success of GTAs' students when compared to tenure track/ tenured faculty's students and other full or part time faculty's students. Specifically, GTAs' students switch STEM intention at a significantly higher percentage than both other types of instructors', and their students lose confidence and interest in mathematics at heightened frequencies when compared to both other instructor types.

Table 5. Student success by instructor type.

Measure of student success:	**Tenure track/ Tenured faculty**	**Other full or part time faculty**	**Graduate teaching assistant**
Percentage of STEM intending students who decided not to pursue Calculus II***	13.9%	9.8%	20.1%
Percentage of students expecting to pass.	96%	95.9%	96.9%
Student change in confidence	-.389 (1.06)	-.440 (1.12)	-.515 (.961)
Student change in enjoyment**	-.255 (1.06)	-.356 (1.09)	-.419 (1.12)
This class has increased my interest in taking more	3.94 (1.40)	3.72 (1.42)	3.58 (1.40)

mathematics.***

Note.* = p $\leq$ .10, ** = p $\leq$ .05, *** = p $\leq$ .001; Standard deviation in parentheses.

Discussion

The above results indicate that in many ways, GTAs are different than other types of Calculus I instructors. They express different beliefs regarding the role of calculators, are more optimistic about their students' capabilities, and are less interested in teaching calculus than tenured/ tenure track faculty and other types of full and part time faculty. Additionally, GTAs report different classroom environments than other types of faculty: students working together more, holding less whole class discussions, and having students give more presentations. While these results indicate some differences between GTAs and other instructor types regarding their beliefs and instructional practices, the most striking differences between GTAs and other instructors lies in their students' success. The students of GTAs decide to not take Calculus II after originally intending to do at much higher frequencies and lose significantly more confidence and interest in mathematics than the students of other instructor types.

These results point clearly to a need to prepare GTAs adequately for the teaching of calculus but also for further examination of the nature and implications of the differences between GTA and other instructor types' beliefs about teaching and teaching practices. Why do GTAs hold a procedural perspective on the role of calculators in the classroom? How does this affect their teaching, and how can we prepare them to explore the conceptual advantages of calculators? Why do GTAs engage their students in more group work and presentations but less whole class discussions? How is this related to their students' decreased interest in studying calculus?

Beyond these questions examining the connections between the above results and student success, are questions regarding the broader implications to teacher preparation at the post-secondary level. In order to teach Calculus at the secondary level in California (a state with some of the most stringent requirements), one must obtain a Bachelor's Degree (or higher) from a credited university, complete a teacher preparation program involving student teaching, and demonstrate subject matter knowledge by passing the California Subject Examinations for Teachers (CSET) or by completing specified mathematics content courses. In order to teach Calculus at the post-secondary level, one must obtain a Bachelor's Degree and be enrolled in a graduate program at the institution, obtain a Master's Degree and teach as an adjunct or obtain a Doctorate and teach as a professor. The difference between these requirements is attention to pedagogical training, which demonstrates differing assumptions on what knowledge is needed to teach mathematics: at the secondary level, content knowledge, pedagogical knowledge, and often pedagogical content knowledge are all prerequisites; at the post-secondary level only strong content level is deemed as sufficient to teach.

Due to this implicit assumption, often the only form of training an instructor receives is as a Graduate student Teaching Assistant (GTA). As such, the training of GTAs is one of few ways to alter the way post-secondary mathematics is taught, and thus the nature and emphases of these training programs are of high significance to the future landscape of post-secondary mathematics. The work described here is the beginning of a large project seeking to respond to the questions outlined above, as well as build a model for GTA training programs that can be used for development of new programs and evaluation of existing programs.

References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? Journal of Teacher Education, 59, 389-407.

Blair, R., Kirkman, E.E., & Maxwell, J.W. (2012), *Statistical abstract of undergraduate programs in the mathematical sciences in the United States*. *Conference Board of the Mathematical Sciences.* American Mathematical Society, Providence, RI.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. Journal for Research in Mathematics Education, 39(4), 372-400

Speer, N., Gutmann, T., & Murphy, T. J. (2005). Mathematics teaching assistant preparation and development. College Teaching, 53(2), 75-80.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15(4), 4–14.

Sowder, J. T. (2007). The mathematical education and development of teachers. In F. K. Lester, Jr. (Ed.), Second handbook of research on mathematics teaching and learning, (pp. 157-223). Charlotte, NC: Information Age Publishers and National Council of Teachers of Mathematics.

Rasmussen, C., & Ellis, J. (2013). Who is switching out of calculus and why? In Lindmeier, A. M. & Heinze, A. (Eds.). *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 4 (pp. 73-80). Kiel, Germany: PME.

GRADUATE STUDENTS' INTEGRATED MATHEMATICS AND SCIENCE KNOWLEDGE FOR TEACHING

Shahram Shawn Firouzian
University of Maine

Previous studies have indicated that effective mathematics teaching relies on teachers' knowledge of both student thinking and mathematical content. Very little is known about the integration (combination) of teacher's mathematical knowledge and science knowledge for teaching important topics like derivative and applied derivative problems. The goal of this study is to examine the knowledge of mathematics and science that teachers draw on when teaching the concept of derivative and applied derivative problems. We conducted task-based interviews with nine graduate assistants (GTAs). Findings revealed that GTAs made use of their knowledge of science as well as of mathematics when discussing how to teach applied derivative problem. In this proposal, we only look at the results of two interviews and try to shed light into the nature of science and mathematics knowledge the teachers use for Teaching and how that can lead into opportunities in professional development for the novice teachers.

Keywords: Teacher Knowledge, Mathematical Knowledge for Teaching, Scientific Knowledge for Teaching, Derivative

Introduction

The knowledge teachers have of mathematics influences how they teach. It appears, however, that additional kinds of knowledge also play important roles in the learning opportunities teachers create for students. The mathematics education community has become increasingly interested in this kind of knowledge and the roles such knowledge plays in teachers' instructional practices (Ball, Lubienski, & Mewborn, 2001; Ball & Bass, 2000; Grossman, 1990; Grossman, Wilson, & Shulman, 1989; Shulman, 1986). This interest includes discussions about how teachers acquire knowledge of student thinking and how preparation and professional development programs can best support teachers' development of such knowledge (Ferrini-Mundy, Burrill, Floden, & Sandow, 2003; Hill, Rowan, & Ball, 2005; Hill, Schilling, & Ball, 2004).

Researchers have documented that elementary and secondary school teachers with richer knowledge of typical student difficulties and strategies create richer learning opportunities for their students (Carpenter, Fennema, Peterson, & Carey, 1988; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989). This research area has served as the basis for professional development programs and associated research projects demonstrating teachers' capabilities in gaining knowledge of student thinking and consequently transforming their instructional practices (Franke, Carpenter, Fennema, Ansell, & Behrend, 1998).

In science education, researchers have also witnessed that students' achievement depends on teachers' pedagogical content knowledge and subject matter knowledge of the science they teach (Magnusson, Krajcik, & Borko, 1999; Davis, & Smithey, 2009; Luft, 2012). In practice, many mathematical ideas are taught using science contexts. However, research into the role of the two domains of science and mathematics knowledge in teaching have been independent of each other. This is the case even though some science education researchers have borrowed the original idea of teachers' "knowledge for teaching" from the mathematics education community. In this present research, we explore how graduate students use their science and mathematics knowledge in their teaching practices.

Although this is part of a larger study focused on novice and experienced mathematics and physics teachers' knowledge for teaching derivative and applied derivative problems, we focus here on findings from only two graduate students' task-based interviews. The analysis shows that they utilize both their science knowledge and their mathematics knowledge while doing the teaching tasks and do so in a way that we refer to as "integrated knowledge." The word integrated is used to describe the bringing together or incorporating of different parts of a teacher's complex system of knowledge for teaching. Our research questions addressing these issues are:

- Can the existing model of teachers' mathematical knowledge for teaching describe teachers' knowledge for teaching applied derivative problems?
- If not, what is the nature of the other knowledge they draw on while teaching these problems?

Theoretical Framework

Researchers have identified multiple domains of mathematical knowledge that teachers use in teaching (McCrory, Ferrini-Mundy, Floden, Reckase, & Senk, 2010; Gess-Newsome, 2002; Magnusson, Krajcik & Borko, 1999; Grossman, 1990; Shulman, 1986). Although the specific boundaries and names of categories vary across publications, in this research we use one of the most agreed upon sets of categories: Ball, Thames, and Phelps's (2008) model (Figure 1) which was produced by modifying the original categories defined by Shulman (1986) in order to more completely describe the knowledge teachers use in teaching mathematics.

<table>
<tr><td colspan="2">Subject Matter Knowledge (SMK)</td><td colspan="2">Pedagogical Content Knowledge (PCK)</td></tr>
<tr><td>Common Content Knowledge (CCK)</td><td rowspan="2">Specialized Content Knowledge (SCK)</td><td>Knowledge of Content and Student(KCS)</td><td rowspan="2">Knowledge of Content and Curriculum (KCC)</td></tr>
<tr><td>Knowledge at the Mathematical Horizon (KMH)</td><td>Knowledge of Content and Teaching (KCT)</td></tr>
</table>

Figure 1. Existing Categories of Mathematical Knowledge for Teaching

Previous studies have investigated students' multiple ways of thinking about derivative and their difficulties in solving applied derivative problems (graphical, optimization and related rate problem). The tasks used in the interviews in this study were inspired by those from existing findings. For instance, White & Mitchelmore (1996) looked at students' difficulties with applied derivative problems like a task which was used in the interviews (shown in Figure 2). The findings about students' difficulties and thinking about derivative were also used from several different sources (Abboud, M., & Habre, S., 2006; Kendal, M., & Stacey, K., 2003; Zandieh, M. J., 2000; Bezuidenhout, J., 1998; White, P., & Mitchelmore, M., 1996; Monk, S., & Nemirovsky, R. (1994); Monk, S., 1994, 1987).

Research Design

Participants and Setting

We interviewed nine graduate teaching assistants (GTAs), six teaching differential calculus, three with more than four semesters experiences and three with no teaching experience.

We also interviewed three GTAs who were teaching physics (classical mechanics) with more than four semesters teaching assistantship experience in introductory physics. All of the interviewees volunteered from a northeastern university for a fifty-minute semi-structured interview. In this paper, we discuss the data from two differential calculus GTAs with more than two years experiences, Viky, and Shai (pseudonyms). They were selected because they both several years of experience with differential calculus classes.

Instrument

As mentioned earlier, our interview tasks were based on previous and ongoing work on students' difficulties and understanding of derivative and applied problems. Using these tasks, we designed an interview protocol (similar to Frank, B., & Speer, N., 2011, 2012) to target the possible domains of knowledge teachers may use for teaching derivative and applied derivative problems. Figure 2 shows one of the tasks used in the interviews.

> *If the edge of a contracting cube is decreasing at a rate of 2 centimeters per minute, at what rate is the volume contracting when the volume of the cube is 64 cubic centimeters? (Provide an explanation for your answer.)*

Figure 2. One of the tasks that interviewees were asked to do and talk about students' difficulties and ideas.

The interviews were semi-structured clinical interviews (Hunting, 1997). The interview had two parts: a section about mathematical knowledge for teaching derivative and a section about their mathematical knowledge for teaching applied derivative problems. The interview questions for both the concepts of derivative and applied derivative problems included:

- How would the participants teach or present the concepts to the students?
- What do they know about students' difficulties with the concept?
- What evidence would they use in assessing the students' works?
- How would they examine samples of students' written work?

Data Analysis Methods

To analyze the interview transcripts, the findings from previous research work on students' understanding of the derivative and applied derivative problems were used. To document the teachers' knowledge for teaching, the domains in Figure 3 were borrowed from existing research work. The acronyms are defined in Figure 1. The interviews were transcribed and the teachers' knowledge of students' thinking and difficulties about derivative were compared to that found in existing research and described using the existing domains of teachers' mathematical knowledge (Figure 3). Analysis included identifying possible domains that they used in explaining or answering the interview tasks.

- SMK: Teachers' abilities to not only understand that something is so but also to understand why it is so.
- SCK: The mathematical knowledge and skills unique to teaching. Knowledge not typically needed for purposes other than teaching.
- PCK: This represents "the blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to diverse interests and abilities of learners, and presented for instruction" (Shulman, 1986, p8).
- KCS: Knowledge that combines knowing about students and knowing about mathematics (Ball, Thames, & Phelps, 2008, p.9).

- KCT: Knowledge that combines knowing about teaching and knowing about mathematics.
- KCC: "Represented by the full range of programs designed for the teaching of particular subjects and topics at a given level, the variety of instructional materials available in relation to these programs, and the set of characteristics that serve as both the indications and contraindications for the use of particular curriculum or program materials in particular circumstances" (Ball, Thames, & Phelps, 2008, p.2).

Figure 3: Descriptions of the Teachers' Knowledge Domains use in the Analysis

Findings

We found that the existing model (Figure 1) did not capture all the elements of the GTAs' mathematical knowledge for teaching. This was particularly evident in the GTAs' responses to the tasks about derivative applications and students' difficulties with this topic. We found that participants were using abilities, skills, knowledge, etc. that were accessed from their domains of knowledge for teaching science. For example, when Viky was asked to give an example of application of derivative, she said:

> I think a good way to start would come if you have a distance graph, showing that you take the derivative to find the velocity at any given point. Oh you know start with the maximums and minimums and when they occur. In general for graphs is to apply those I guess more in real world situation.

Using the existing model of teacher's mathematical knowledge for teaching, we propose that Viky is using her SMK since she is using a physics example as an applied problem but also knows the reason for using it being "more in real world". We also propose that she is using her KCS and KCC elements of PCK since distance-velocity applied derivative problems are a particular way of introducing derivative applications to the students. However, it seems that Vicky's use of the concepts of velocity and distance from physics shows she is accessing knowledge from beyond just mathematic and is using integrated mathematics and science knowledge for teaching applied derivative ideas.

As a second example, Shai was asked to explain why someone who is not a mathematics major should care about derivative or if he could give examples of where the concept of derivative might be used. He responded: "if [he is a] physics student, it is very important for him because he always needs to deal with velocity, acceleration." Like Viky, it seems Shai is accessing KCC and KCS elements of PCK to give a particular example of applied derivative problem but he continues saying "I always think biology has a lot of relationship with mathematics because like well, I think they don't need to use any math when they are undergrad but when they want to go to graduate school, math becomes very important for them." Again it seems as though he is accessing his mathematics and science knowledge for teaching applied derivative problems within the KCC and KCS sub-categories of PCK.

As was the case with Vicky, the broad categories of KCC and KCS do not account for Shai's knowledge of examples of other disciplines. One can argue that the science knowledge he has about application of derivative in other discipline can be categorized as an expansion of PCK within KCC and KCS categories and perhaps this is where it overlay into the teacher's science knowledge for teaching.

Conclusions & Implications

The previous examples as well as additional analysis of both participants' interviews are consistent with the claim that knowledge is complex and cannot be easily categorized into distinct and separate domains. By using the existing model, we noticed that participants were using two domains of knowledge simultaneously. The term integrated were used because our analysis showed the complex nature of teachers' knowledge especially the complex nature of integrated mathematical and scientific knowledge for teaching applied derivative problems. These are preliminary results, and we expect further analysis to reveal the nature of integrated mathematical and scientific knowledge for teaching.

As we see from Viky's and Shai's explanations, the mathematical knowledge required for teaching is indeed multidimensional therefore the professional education should perhaps be organized to help teachers learn the range of knowledge and skills they need in teaching. By this research we are hoping to identify the opportunities that we can provide our teachers with, in order to learn mathematics and sciences necessary for teaching mathematics and propose development of effective professional developments.

Discussion Questions

- Can this category of knowledge be explained in the context of knowledge of content and curriculum?
- What are the possible approaches we can adopt into our professional development of the novice teachers that can enhance their scientific and mathematical knowledge for teaching?

References

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics*. Westport, CT: Ablex.

Ball, D. L., Lubienski, S., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of Research on Teaching* (pp. 433-456). Washington, DC: American Educational Research Association.

Ball, D. L., Thames, M. K., & Phelps, G. (2008). *Content Knowledge for Teaching*, Journal of Teacher Education 59, 389-407.

Bezuidenhout, J. (1998). First-year university students' understanding of rate of change. *International Journal of Mathematical Education in Science and Technology*, *29*(3), 389-399.

Carpenter, T., Fennema, E., Peterson, P., & Carey, D. (1988). Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education, 19*, 385-401.

Carpenter, T., Fennema, E., Peterson, P., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*(4), 499-531.

Davis, E. A., & Smithey, J. (2009). Beginning teachers moving toward effective elementary science teaching. *Science Education*, *93*(4), 745-770.

Ferrini-Mundy, J., Burrill, G., Floden, R., & Sandow, D. (2003). *Teacher knowledge for teaching school algebra: Challenges in developing an analytical framework.* Paper presented at the Annual meeting of the American Educational Research Association, Chicago, IL.

Frank, B. & Speer, N. (2011). On the Job Learning: Instructors' Development of Knowledge for Teaching. *Presentation at No Question Left Behind: Bringing Guided-Inquiry Curricula into Science and Mathematics Classrooms conference.* Orono, ME.

Frank, B., & Speer, N. (2012). On the Job Learning: Instructors' Development of Knowledge for Teaching. *In preparation for submission to Cognition & Instruction*.

Franke, M., Carpenter, T., Fennema, E., Ansell, E., & Behrend, J. (1998). Understanding teachers' selfsustaining, generative change in the context of professional development. *Teaching and teacher education, 14*, 1-14.

Gess-Newsome, J. (1999). *Pedagogical content knowledge an introduction and orientation.* In J. Gess-Newsome & N. G. Lederman (Eds.), Examining pedagogical content knowledge (pp. 3–20). Dordrecht: Kluwer.

Gess-Newsome, J. (2002). Pedagogical content knowledge: An introduction and orientation. In *Examining pedagogical content knowledge* (pp. 3-17). Springer Netherlands.

Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.

Grossman, P. L., Wilson, S., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. In M. C. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 23-36). Oxford: Pergamon Press.

Habre, S., & Abboud, M. (2006). Students' conceptual understanding of a function and its derivative in an experimental calculus course. *The Journal of Mathematical Behavior*, *25*(1), 57-72.

Hill, H., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hill, H., Schilling, S., & Ball, D. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal, 105*(1), 11-30.

Hunting, R. (1997). *Clinical Interview Methods in Mathematics Education Research and Practice*. Journal of Mathematical Behavior, 16(2), 145–165.

Kendal, M., & Stacey, K. (2003). Tracing learning of three representations with the differentiation competency framework. *Mathematics Education Research Journal, 15*(1), 22-41.

Luft, Julie. "Subject-specific induction programs: Lessons from science." *NSSE Yearbook* 111.2 (2012): 417-442.

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. Lederman (Eds.), Examining pedagogical content knowledge: The construct and its implications for science education (pp. 95 – 132). Dordrecht, The Netherlands: Kluwer

McCrory, R., Floden, R., Ferrini-Mundy, J., Reckase, M., & Senk, S. (2010). Knowledge *of algebra for teaching: Framework of knowledge and practices*. Journal for Research in Mathematics Education, 43(5), 584-615.

Monk, S. (1987). Students' understanding of function in calculus courses. *Unpublished manuscript.*

Monk, S. (1994). Students' understanding of functions in calculus courses. *Humanistic Mathematics Network journal, 27,* 21-27.

Monk, S., & Nemirovsky, R. (1994). The Case of Dan: Student Construction of a Functional Situation through Visual Attributes. In E. Dubinsky, A. H. Schoenfeld & J. Kaput (Eds.), *Research in Collegiate Mathematics Education I* (Vol. 4, pp. 139-168). Providence, RI: American Mathematical Society.

Shulman, L. S. (1986). *Those who understand: Knowledge growth in teaching*. Educational Researcher, 15(2), 4–14.

White, P., & Mitchelmore, M. (1996). Conceptual knowledge in introductory calculus. *Journal for Research in Mathematics Education*, 79-95.

Zandieh, M. J. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. *Research in Collegiate Mathematics Education. IV. CBMS Issues in Mathematics Education*, 103-127.

Working Together on Mathematics Homework:
A Look at How University Students Spend Their Time Outside the Classroom

Gillian Galle
New York University

Despite the large amount of time university students are expected to spend studying material and learning on their own outside of the classroom, little is known about what specific student study habits look like. This study sought to start developing a description of what activities students engage in when studying together in self-formed groups outside of the classroom. By identifying a set of macrotasks, verbally-cued transactions that identify what activity the group is currently engaged in doing, this study provides a way to compare how different study groups allocate their time and distinguish between the enactment of social and sociomathematical norms outside of the classroom.

Keywords: Group Work, Study Habits, Norms

Introduction

There's an expectation that university students spend up to 3 times the number of hours spent in class working on their own (Wu, 1999). While several studies address student study habits, they do not provide much information on how students spend their time studying outside of the classroom. At best they provide confirmation of the ideas that students work with peers outside of the class while studying and exhibit a variety of study behaviors. However, these studies base their conclusions on self-reported data collected via surveys and interviews and do not provide any descriptive detail addressing what these self-reported study behaviors look like in action.

This study sought to add the missing descriptive element of what transpires when students study together in groups by answering the research question: How are students spending their time while working together?

Framework

Several important ideas contribute to how we understand students working together in groups and why we can expect to see patterns in their interactions. Rogoff's (2003) description of guided participation provides some insight into why students negotiate a common understanding while working together. Lave and Wenger (1991) also recognize the creation of shared resources in their communities of practice. One way of interpreting these shared resources and communally developed interactions is by comparing them with Cobb and Yackel's (1996) social and sociomathematical norms.

Guided Participation and Communities of Practice

"Guided participation provides a perspective to help us focus on the varied ways that children learn as they participate in and are guided by the values of and practices of their cultural communities" (Rogoff, 2003, p. 283-284). Thus, guided participation can reasonably account for varying forms of participation in a variety of sociocultural activities such as study groups or communities of practice (Lave & Wenger, 1991).

Rogoff (2003) explains guided participation as the result of coordinating two basic processes: the bridging of meaning and the structuring of participation. The individuals involved have to develop a common language or perspective in order to share their ideas with each other as well as negotiating the ways in which they interact. Peers that share equal levels of understanding would enact these two processes as a collaborative negotiation. In the event that the skill level varies between participants, the decisions made by the expert peer, or the peer that is more confident with the material, will more heavily influence the negotiation that occurs during guided participation.

Similar to Rogoff's guided participation, Lave and Wenger (1991) defined legitimate peripheral participation in communities of practice. Communities of practice are comprised of three dimensions (Wenger, 1998, para. 8):

(1) **What it is about**—its *joint enterprise* as understood and continually renegotiated by its members
(2) **How it functions**—the relationships of *mutual engagement* that bind members together into a social entity
(3) **What capability it has produced**—the *shared repertoire* of communal resources (routines, sensibilities, artifacts, vocabulary, styles, etc.) that members have developed over time.

Due to its fluid definition, communities of practice can be found anywhere. Within the context of this study then, there is a community of practice at the university level, at the classroom level, and at the study group level. The first two dimensions of community of practice at the level of the student study groups become:

(1) **What it is about** – students gathering in order to perform activities related to furthering their class preparation
(2) **How it functions** – the tasks, especially those that are frequently repeated, that the students participate in and what the students bring with them to contribute to the group's overall endeavors

The third dimension, the shared repertoire developed within the study group, will be comprised of all the elements Wenger listed. Of particular interest to this study are the routines and repeated interactions that give structure to the time the study groups spend working together. While many of these interactions are governed by social and cultural cues, when students are working on mathematics problems together they will also be utilizing implicit rules for how they propose new ideas and problem solving strategies and how they defend these ideas to their peers. These shared activities that comprise the group's shared repertoire of communal resources and may also be included in what Yackel and Cobb (1996) term social and sociomathematical norms.

Social and Sociomathematical Norms

Cobb and Yackel (1996) state that classroom social norms "characterize regularities in communal or collective classroom activity and are considered to be jointly established by the teachers and students as members of the classroom community" (p. 178). These norms are not specific to the content of the class, so classroom social norms such as *explaining an answer* could be found in a history classroom as well as a mathematics classroom.

Subsequently, Cobb and Yackel identified sociomathematical norms as the "normative aspects of whole-class discussions that are specific to student's mathematical activities" (Cobb & Yackel, p. 178). So while typical classroom social norms include the expectation that students will engage in the explanations, justifications, and argumentation, sociomathematical norms require students to understand mathematical difference, mathematical sophistication, and acceptable mathematical explanation and justification (Yackel & Cobb, 1996, p. 461).

Although Cobb and Yackel's (1996) interpretation of the social perspective was meant to provide a way to describe the activities enacted within a classroom, it has implications for any community of practice. In particular, it may be applied to student study groups since the study groups that are formed by students outside of the classroom constitute their own communities of practice with their own cultural expectations.

Data Collection and Analysis

Participants were all students of at least second year standing that were enrolled in an undergraduate course that blended topics from linear algebra, differential equations and multidimensional calculus. While the instructor of this course encouraged students to collaborate in class, none of the groups from this class that worked together over the semester, either inside or outside of class time, were the result of assignment by the instructor or the researcher. All groups were self-formed and self-directed.

Students were observed studying together in groups outside of the classroom in a space that was equipped with tables, chairs, internet access, and white boards. Video-recordings were made of all study sessions and were supplemented with field notes and journal entries that students completed at the end of each session.

All study sessions were transcribed and coded in several passes using a combination of Goos, Galbraith, and Renshaw's (2002) coding scheme and Blanton, Stylianou, and David's (2009) coding schemes. While the combination of these two sets of codes encompasses most of the interactions that could be expected, the researcher also employed an open coding scheme in order to add codes as needed should an utterance defy categorization in either of the two schemes.

This paper focuses on the results of analyzing two study groups as they worked on the same homework assignment over the course of a week. Group A was comprised of Abigail, Josh, Amy, and Zoey while Group B was made up of Hugh, Phil, and Ben. Names have been changed to protect the anonymity of the participants. The assignment they were working on was given during the middle of the semester, thus patterns in behavior and routines could be taken as established norms.

Results and Significance

By performing a discourse analysis several times and varying the grain size under scrutiny, I identified a set of macrotasks, or general activities that the students of both groups engaged in over the course of a study session. While many of these macrotasks are general enough to be occurrences in study group sessions dedicate to non-mathematical work, similar to social norms, several of them were avenues for observing course-specific interactions, or sociomathematical norms.

Macrotasks

Throughout a study session, student dialogue gives insight into what objectives the group, or the individual participating in the group, is currently focused on. By reviewing the discourse analysis, I was able to identify several themes that arose repeatedly. I refer to them as macrotasks to distinguish them from the specific steps that may be taken to achieve them. For instance, *doing problems* would be considered a macrotask while actions such as *reading the problem* or *calculating a solution* would be tasks that the students need to complete within that macrotask. Thus we can view a study session as being comprised of a series of macrotasks,

The majority of each observed session was dedicated to the macrotask of *doing problems*. As its name suggests, portions of the session receiving this designation featured the students reading problems, discussing methods of solution, and comparing answers. Other macrotasks that arose throughout the sessions included:

- *Getting situated* – covers interactions such as greeting each other upon arrival, physically re-arranging tables, chairs, and seating arrangements, procuring necessary materials like pencils, paper, or calculators.
- *Assignment (HW) planning* – Discussions regarding what problem is going to be worked on next.
- *Session planning* – Negotiations over how long the group will be able to work that day, what goals the individuals would like to accomplish during this particular session.
- *Checking the group's progress* – Typically enacted whenever an individual joins the group study session, it covers a comparison of what problems the group does and does not have solutions for.
- *Planning to ask the professor a question* – Discussions revolving around what questions to ask the professor, typically during office hours, and negotiations over which individual will be responsible for visiting the professor to ask the questions.
- *Reporting the professor's response* – Discussions in which an individual relays the professor's response to the group's questions.
- *Discussing the homework write-up* – Discussions relating to proper presentation of a problem's solution such as which steps to include and what comprises a complete argument.
- *Planning future study sessions* – Negotiations over date, time, and location for future meetings of the study group.
- *Off-topic discourse* – Any dialogue that does not directly relate to the objectives for the study session, the course content, or the class sessions themselves.
- *Recognizing off-topic talk* – Acknowledgement that an off-topic conversation has taken the group off course or has gone on for too long accompanied by efforts to refocus the group on the task or problem at hand.
- *Filling out journal entries* – Specific to the nature in which the data was collected, there was a developed norm of students completing their journal entries and checking with each other regarding amount of time they spent in the study space for the particular session and which problems they felt they had completed.

As can be seen in Table 1, both groups exhibited many of these macrotasks over the two observed study sessions included in this analysis. In both groups the majority of their time was spent *doing problems* with *off-topic discourse* taking up the second largest amount of time. Group A and Group B also spent similar amounts of time getting situated. However, there were also differences in the macrotasks enacted by each group. Group A spent more time on *checking the group's status* and *session planning* while Group B spent more time on *assignment (HW) planning*.

Additionally, as the two groups tended to meet at different times of day they had different levels of access to the professor. Since Group A typically met in the mid-afternoon during the professor's office hours, whereas Group B met in the evenings, Group A spent more time on the macrotasks of planning to ask the professor a question and reporting the professor's response.

Macrotask	Macrotask Enactment Over Observed Sessions: Minutes (%)	
	Group A (230.45 Minutes Total)	Group B (231.532 Minutes Total)
Doing Problems	127.166 (55.18%)	161.300 (69.67%)
Getting Situated	6.067 (2.63%)	5.466 (2.36%)
Checking the Group's Status	5.717 (2.48%)	2.516 (1.09%)
Assignment (HW) Planning	3.916 (1.70%)	6.884 (2.97%)
Session Planning	4.583 (1.99%)	1.366 (0.59%)
Off-Topic Discourse	72.917 (31.64%)	41.950 (18.12%)
Recognizing Off-Topic Talk	0 (0.00%)	0.550 (0.24%)
Planning to Ask Prof. a Question	1.684 (0.73%)	0 (0.00%)
Reporting Professor Response	1.983 (0.86%)	0 (0.00%)
Homework Write-Up	1.617 (0.70%)	0 (0.00%)
Future Study Sessions	1.783 (0.77%)	0 (0.00%)
Journal Entries	3.017 (1.31%)	11.500 (4.97%)

Table 1. A comparison of amount of time spent on each macrotask by Group A and Group B.

Enactment of Normative Behaviors

Yackel and Cobb claim that the "social norms implicit in the inquiry approach to mathematics instruction […] foster [students'] development of social autonomy" and that sociomathematical norms "foster[s] the development of intellectual autonomy" (1996, p.473). Thus the macrotasks identified in this study provide a way to look for enactments of the social and sociomathematical norms in student behavior observed outside of the classroom context, where the students practice their social and intellectual autonomy.

<table>
<tr><th></th><th colspan="2">Norms</th></tr>
<tr><th></th><th>Social</th><th>Sociomathematical</th></tr>
<tr><td rowspan="8">Macrotasks</td><td>Getting Situated</td><td rowspan="2">Doing Problems</td></tr>
<tr><td>Assignment (HW) Planning</td></tr>
<tr><td>Session Planning</td><td rowspan="2">Planning to Ask the Professor a Question</td></tr>
<tr><td>Checking the Group's Progress</td></tr>
<tr><td>Planning Future Study Sessions</td><td rowspan="2">Reporting the Professor's Response</td></tr>
<tr><td>Off-Topic Discourse</td></tr>
<tr><td>Recognizing Off-Topic Talk</td><td rowspan="2">Discussing the Homework Write-up</td></tr>
<tr><td>Filling Out Journal Entries</td></tr>
</table>

Table 2. Classification of the macrotasks across norms.

Macrotasks such as *doing problems*, *deciding to ask the professor*, and *reporting the professor's response* all featured moments of students enacting sociomathematical norms as negotiated by their group. Other macrotasks, such as *checking the group's progress* and *getting situated*, are strictly social. Thus the norms the students enact within this latter set of macrotasks would also be considered socially established routines and may be generalizable to study groups for other courses with non-mathematical content. Table 2 provides a complete classification of the macrotasks with respect to the norms students may enact during the task.

Conclusions and Directions for Future Study

The significance of the macrotasks is two-fold. First, the identification of macrotasks helps give an overview of what transpires during a study session and provides a way to compare what different study groups are spending time on. In this respect, the results of this study mark one of the first efforts in the field to provide a meaningful method for describing what occurs in student study groups outside of the classroom.

The other contribution made by the identification of macrotasks is the start of identifying social and sociomathematical norms outside of the classroom environment. Given the hypothesis that student behavior is learned in the classroom, especially during group work activities, and can be influenced by behaviors modeled by the instructor (Yackel & Cobb, 1996), a basis for comparison with student behaviors outside of the classroom is needed. Left to their own devices, the students in this study often resorted to the methods of verification that they are most comfortable with, typically a very procedural approach to the verification of their answers to problems despite the inquiry-based techniques their professor promoted.

In light of this, the findings of this study lay the groundwork for future studies in what sort of transfer exists from in-class student behavior to student behavior out of the classroom. It may be that if an instructor wants his students to engage in more conceptual dialogue outside of the classroom he may have to do more than implicitly demonstrate the behavior in class. It could be case that the instructor needs to explicitly draw the students' attention to the activities and discourse demonstrated in class in order to introduce changes in the students' behaviors outside of the class.

References

Blanton, M., Stylianou, D., & David, M. (2009). Understanding instructional scaffolding in classroom discourse on proof. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective*, (pp. 290-306). New York, NY: Routledge.

Cobb, P. & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, *31*(3/4), 175-190.

Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics*, *49*, 193-223.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation (Learning in doing: Social, cognitive, and computational perspectives)*. New York, NY: Cambridge University Press.

Rogoff, B. (2003). *Cultural nature of human development*. Cary, NC: Oxford University Press, Inc.

Wenger, E. (1998). Communities of practice. Learning as a social system. *Systems Thinker*. Retrieved from: http://www.co-i-l.com/coil/knowledge-garden/cop/lss.shtml

Wu, H. (1999). The joy of lecturing – With a critique of the romantic tradition in education writing. In S. G. Krantz (Ed.), *How to teach mathematics* (2nd ed.), (pp. 261-271). Providence, RI: American Mathematical Society.

Yackel, E. & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4), 458 – 477.

MATHEMATICS TEACHERS' MODLES OF QUANTITATIVE REASONING

David Glassmeyer, Kennesaw State University
Michael Oehrtman, University of Northern Colorado
Jodie Novak, University of Northern Colorado

The purpose of this study was to document mathematics teachers' models of quantitative reasoning as they participated in a Model Eliciting Activity (MEA) grounded in their classroom practice. This MEA was designed and implemented in a master's course of 21 in-service mathematics teachers. The documents produced by the teachers were analyzed using a models and modeling perspective to determine how teachers' models of quantitative reasoning developed through the MEA. Findings from this study included a framework describing the two ways teachers developed their model of quantitative reasoning. First, teachers' models of quantitative reasoning became more coherent by being better articulated and connected between abstract and practical settings. Second, the middle school teachers' models became more aligned with mathematics education literature by referring to quantities and quantitative relationships as aspects of quantitative reasoning, while most high school teachers' models did not become more aligned with literature.

Key words: quantitative reasoning, mathematics teacher education, model-eliciting activity

Introduction and Literature Review

Mathematics education literature lacks details on how teachers think about mathematical ideas presented in education reform efforts, especially the Common Core State Standards for Mathematics (Confrey & Krupa, 2010; Sztajn, Marrongelle, & Smith, 2011). One focus for research efforts has been to investigate and promote mathematics teachers' quantitative reasoning (Thompson, 2011, 2013). This purpose of this study was to address this focus by answering the research question: How do mathematics teachers' models of quantitative reasoning develop through a Model Eliciting Activity (MEA) grounded in their classroom practice?

To frame the study, we first clarify definitions regarding quantitative reasoning and MEAs for teachers. The work of Thompson (1990, 2011) and colleagues (Smith & Thompson, 2008) offer a theory of quantitative reasoning, highlighting the role of learners constructing quantities and quantitative relationships. Quantities are a cognitive object, according to Thompson's theory, and are composed of four components: (a) an object, event, or idea, (b) a measureable attribute, (c) a unit of measurement, and (d) a conceivable numerical value or possible values. Quantitative relationships are formed as a person's conceptualized quantities are joined in quantitative operations. Thompson (1990, 2011) calls a person's mental network of quantities and quantitative relationships their quantitative structure. This structure may contain multiple layers, all within the individual's mind rather than in the world. Thompson views quantitative reasoning as the mental process where a person's quantitative structure is used to achieve a goal. Quantitative relationships differ from numerical relationships, which deal only with arithmetic operations. Moore, Carlson, and Oehrtman (2009) summarize quantitative reasoning, in light of Thompson's theory, as attending to and identifying quantities, identifying and representing quantitative relationships, and constructing new quantities. In the remainder of this document, we refer to both Thompson (1990, 2011) and Moore et al.'s (2009) definitions of quantity, quantitative relationship, and quantitative reasoning for a common reference point for what is meant by the term.

Teacher MEAs provide a way to investigate teachers' ways of thinking about mathematics and their practice. MEAs are tasks that engage teachers in thinking about realistic and complex problems embedded in their practice in order to foster ways of thinking that can be used to communicate and make sense of these situations (Doerr & Lesh, 2003; Lesh & Zawojewski, 2007). MEAs have been shown to contribute to teacher development because these activities make teachers engage in applicable mathematics, consider student reasoning more deeply, and reflect on beliefs about problem solving (Chamberlin, Farmer, & Novak, 2008; Schorr & Koellner-Clark, 2003; Schorr & Lesh, 2003). While these studies have implemented successful MEAs for teachers, there is a need for additional activities given the recent demands the CCSSM place on teacher education programs (Confrey & Krupa, 2010; Garfunkel, Reys, Fey, Robinson, & Mark, 2011).

Methods

The theoretical perspective we used for the study is a Models and Modeling Perspective, as described by Lesh and colleagues. A Models and Modeling Perspective incorporates MEAs to simultaneously investigate and improve teachers' models, or systems of interpretation, within educational problem-solving situations. Thus this perspective provided us a powerful lens for understanding teachers' ways of thinking, their development, and provided a mechanism for analyzing and piecing together findings (Koellner-Clark & Lesh, 2003; Hiebert & Grouws, 2007; Sriraman & English, 2010).

For this study, we designed and implemented an MEA in a master's course of 21 in-service mathematics teachers. We focused the study on a newly developed mathematics education course called Quantitative Reasoning in Secondary Mathematics, which was offered four weeks in the summer. The MEA, worth 50% of the course grade, asked teachers to create and refine a quantitative reasoning task for their students with the intention of teachers implementing the task in the following fall. Teachers worked in groups of three or four and were roughly clustered by the content they taught: the high school Groups 1, 4, and 5 taught algebra 2, pre-calculus, and trigonometry, while the middle school Groups 2, 3, and 6 taught algebra 1 or pre-algebra. Each group received MEA feedback from several sources, including the instructor, each other, undergraduate students, and in some cases, their own students. Each type of feedback prompted an updated iteration of the task and supporting documents that captured how teachers' models develop. Data collection consisted of the iterations of documents generated by the MEA (see Table 1). The Pre-Assignment and Version 5 were individual documents, while Versions 1-4 were group documents. Using content analysis on the documents, we identified patterns in the ways teachers' thinking about quantitative reasoning tasks developed due to this process.

Table 1. Summary of Quantitative Reasoning (QR) documents analyzed

Assignment Name	**Short Description of Components**
Pre-Assignment	Document including initial models of QR, QR tasks, QR course
Version 1	Four documents including (a) Quantitative Reasoning Task; (b) Facilitator Instructions; (c) Assessment Guidelines; (d) Decision Log
Instructor's Feedback	Instructor's comments and suggestions to Version 1.
Version 2	Updated Version 1 in response to the instructor's feedback.
Teachers' Feedback	Groups swap Version 2 and offer comments and suggestions
Version 3	Updated Version 2 in response to the teachers' feedback.
Undergraduate Work	Student work after completing QR task (part (a) of Version 3)
Version 4	Updated Version 3 in response to student work, plus evaluation of

	student work.
Version 5	Updated Version 4 in response to the feedback received from implementation, plus the actual evaluation of the K-12 student work.

Findings

We found two basic structures comprised the framework about mathematics teachers' models of quantitative reasoning. First, teachers' models of quantitative reasoning became more coherent by being better articulated and connected between abstract and practical settings. Specifically, teachers more clearly described aspects of quantitative reasoning and referred to these aspects when attending to their classroom practice. Second, as the MEA progressed, the middle school teachers recognized aspects of quantitative reasoning in ways that became more aligned with mathematics education literature. Specifically, by the MEA conclusion middle school teachers made statements referring to quantities and quantitative relationships, while high school teachers tended not to make such statements. Since quantities and quantitative relationships are vital components of quantitative reasoning according to Thompson (2011) and other leading mathematics education researchers (Moore et al., 2009), this collection of statements indicated middle school teachers' models became more aligned with literature. This section details each basic structure underling this framework and includes reasons for changes in teachers' models.

First, teachers better articulated and connected aspects of quantitative reasoning across different settings. We make a distinction between how teachers provided information about quantitative reasoning in abstract settings, which is not specific to any context, and the information teachers provided in practical settings, such as when teachers were designing, implementing, and evaluating actual instructional activities. Teachers initially provided less information regarding quantitative reasoning in abstract settings. For instance, at the start of the MEA three individual teachers and Groups 2, 5, and 6 did not provide information about quantitative reasoning in abstract settings.

The teachers that did initially make statements referring to aspects of quantitative reasoning did so by identifying aspects in abstract settings that were different than the aspects they said in practical settings. An example of one teacher's Pre-Assignment responses that referred to different aspects of quantitative reasoning between abstract and practical settings was Joyce. Joyce abstractly defined quantitative reasoning as being "strongly associated with number sense and the ability to visualize (or conceptualize in some way) certain amounts." When describing quantitative reasoning in practical settings, such as what this looks like in her classroom, she stated,

When I teach lessons, my goal is to help students think quantitatively as we work through problems. I want them to make sense of what they are doing, not to just do it…when my students and I work with logarithms, I spend a lot of time discussing what a particular problem means. In general (overall), I do not constantly give lengthy explanations so as not to cause algebraic processes to become tedious and disjointed, but these explanations are necessary at the appropriate times.

Joyce did not mention number sense or visualization of amounts as aspects of quantitative reasoning in this setting, but instead focused on sense-making when working with functions in a problem. While number sense could have been included in sense-making, Joyce's responses did not provide evidence of this, and thus she did not connect her aspects of quantitative reasoning across her responses in a way we could observe. Similar to these patterns in individual responses, Groups 1, 3, and 4 made statements referring to aspects of quantitative reasoning that differed across abstract and practical settings. Only six teachers initially provided statements about quantitative reasoning that were connected across abstract and practical settings

By the conclusion of the MEA, every teacher provided more information about aspects of quantitative reasoning in abstract settings. All 6 groups and all 21 teachers made statements referring to at least one aspect of quantitative reasoning that was connected across abstract and practical settings. For example, Charlotte recognized the aspects of quantities and quantitative relationships in her final MEA documents, first by saying in her Decision Log that quantitative reasoning was:

making sense of a problem by trying to visualize in your mind a model, interpreting data by breaking it down so one can identify relevant quantities and their meanings, representing relationships between quantities using graphs, tables, and algorithms then trying to create a formula through that reasoning. It's essential for students to focus on recognizing relationships and having them write or explain their thought processes in how quantities relate to one another and showing they work together in a process not individually, as well as, constructing new quantities that are not given to form a conclusion.

Here she identified quantities as an aspect of quantitative reasoning, and referred to quantitative relationships by considering how quantities covary in relationships and how these relationships create new quantities. Thus Charlotte made statements that referred to the aspects of identifying quantities and quantitative relationships in this abstract setting. These aspects are seen in the next paragraph when she referred to the practical setting in the context of her group's task:

Our group's MEA relates to quantitative reasoning when we have students reason about which would be the best fundraiser for their school and explaining why it would be the best choice, identifying quantities (varying and not), determining what quantities mean and how they relate to each other, creating visuals to identify relationships, having students explain what it means to have quantities co-vary, constructing general equations through these discoveries, and presenting their work to peers and teachers.

Since quantities and quantitative relationships were in her group's task, Charlotte's response from this practical setting here shared aspects with her earlier response given in an abstract setting. Thus Charlotte's responses emulated how teachers better articulated and connected aspects of quantitative reasoning across these settings as the MEA progressed. Teacher reflections about their development suggested peer feedback, student feedback, and course materials influenced them to better articulate and connect aspects of quantitative reasoning.

The second basic structure in how teachers' models of quantitative reasoning developed was that the middle school teachers recognized aspects of quantitative reasoning in ways that became more aligned with mathematics education literature. At the beginning of the study, teachers attended to few aspects of quantitative reasoning. Only two teachers made statements suggesting quantities were an aspect of quantitative reasoning, and only one teacher made a statement suggesting quantitative relationships was an aspect of quantitative reasoning. Furthermore only one of the six groups initially made statements in their Version 1 documents referring to quantities and quantitative relationships as aspects of quantitative reasoning. Instead, most teachers made statements about what we called pseudo-quantities in their initial MEA documents. Pseudo-quantities are numerical values, unknowns, or other features of a contextual setting where the teachers did not fully distinguished the object, attribute of the object, and units of the object being considered. For example, Penny gave the response that quantitative reasoning was "giving students a problem involving quantities where they have to determine a strategy for solving the problem," with no further statements about what was meant by "quantities." Since her use of this word was vague and had no evidence of attending to an object, a measurable attribute of the object, a way to assign values to this measure, or an accompanying unit, her response was coded as referring to pseudo-quantities. Rather than quantitative relationships, teachers identified numerical relationships, which are the conceptual relationships between variables or measures of quantities based only on the arithmetic operations involved.

During the MEA, middle school teachers made statements indicating their models of quantitative reasoning became more aligned with mathematics education literature. The three middle school groups (2, 3, 6) all made statements in their final MEA documents that referred to both quantities and quantitative relationships as an aspect of quantitative reasoning. For example, Group 2 incorporated a table "designed to help [students] think critically about what quantities would be present in fundraising situations." This table was in the Quantitative Reasoning Task and had accompanying expectations in the Assessment Guidelines that asked students to identify the object, attribute, unit, for "all of the varying and unvarying quantities that are present in a fundraising situation." These expectations indicated Group 2 referred to quantities as an aspect of quantitative reasoning by the MEA conclusion. Teachers said student feedback prompted them to be more explicit about quantities in their MEA documents. Peer feedback and course materials, such as Pathways to Calculus (Carlson & Oehrtman, 2011), gave teachers examples of how quantities could be incorporated in their MEA. Middle school teachers did not say why they included quantitative relationships, rather than numerical relationships, in their MEA documents.

Most of the high school teachers' models of quantitative reasoning did not develop in ways aligned with literature. These teachers continued to refer to pseudo-quantities and numerical relationships when making statements about quantitative reasoning. The exception to this pattern was Group 5. Group 5 referred to quantities and quantitative relationships as aspects of quantitative reasoning in both initial and final MEA documents. Two of the teachers in Group 5 also made statements about coordinating quantitative relationships at the MEA conclusion.

Implications

This study has three main implications for the field of mathematics education. First, this study supports and extends prior work regarding how teacher MEAs can document teachers' models within teacher education settings. This study provided evidence an MEA can document development in teachers' models and provide answers to research questions regarding mathematics teachers' models of quantitative reasoning. Designing and implementing these successful activities for teachers is both novel and significant to the field of teacher education (Confrey & Krupa, 2010; Garfunkel et al., 2011). Second, the framework created by answering this study's research question offers researchers a tool to better understand in-service teacher thinking about quantitative reasoning beyond the context of this study. Researchers need ways to understand teacher thinking, including how they think about quantitative reasoning and other CCSSM standards for mathematical practice (Confrey & Krupa, 2010; Sztajn et al., 2011; Thompson, 2013). This study was designed to address these needs, and does so by providing a novel framework for how one population of teachers thought about quantitative reasoning.

Third, this study established sharable practices other teacher educators can use to develop teacher ways of thinking about quantitative reasoning that are connected to practice and aligned with literature. Based on this study's findings, we recommend that teacher educators have mathematics teachers develop a task and supporting documents for their own classroom practice, and provide the opportunity for teachers to revise their documents after receiving various forms of feedback. Feedback from students and peers promoted teachers to address inconsistencies when thinking about quantitative reasoning across different settings, while course materials helped teachers align their ways of thinking about quantitative reasoning with literature. Future research should focus on identifying ways to provide extra support for secondary teachers' thinking about quantities and quantitative relationships. Future research can also continue determining how mathematics teachers think about quantitative reasoning and finding ways to support productive ways of thinking.

References

Carlson, M. & Oehrtman, M. (2011). *Precalculus, Pathways to Calculus: A Problem Solving Approach*, First Edition. Phoenix, AZ: Rational Reasoning.

Chamberlin, M., Farmer, J., & Novak, J. (2008). Teachers' perceptions of assessments of their mathematical knowledge in a professional development course. *Journal of Mathematics Teacher Education, 11*(6), 435-457.

Clark Koellner, K., & Lesh, R. (2003). A modeling approach to describe teacher knowledge. In R. Lesh, & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics teaching, learning, and problem solving* (pp. 159–173). Mahwah, NJ: Lawrence Erlbaum Associates.

Confrey, J., & Krupa, E. (2010). Curriculum design, development, and implementation in an era of common core state standards: Summary report of a conference.

Doerr, H. M., & Lesh, R. A. (2003). A modeling perspective on teacher development. In R. A. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning and teaching* (pp. 125-39). Mahwah, NJ: Lawrence Erlbaum Associates.

Garfunkel, S., Hirsch, C., Reys, B., Fey, J., Robinson, R., & Mark, J. (2011). A Summary Report from the Conference "*Moving Forward Together: Curriculum & Assessment and the Common Core State Standards for Mathematics*". Arlington, VA.

Hiebert, J., & Grouws, D. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 371–404). Charlotte, NC: Information Age Publishing.

Lesh, R. & Zawojewski, J. S. (2007). Problem solving and modeling. In F. Lester (Ed.). *The Second Handbook of Research on Mathematics Teaching and Learning* (pp. 763-804). Charlotte, NC: Information Age Publishing.

Moore, K., Carlson, M., & Oehrtman, M. (2009). The Role of Quantitative Reasoning in Solving Applied Precalculus Problems, *Proceedings of the Twelfth Conference on Research in Undergraduate Mathematics Education*, 26 pages, Web publication at http://rume.org/crume2009/Moore1_LONG.pdf

Schorr, R. Y., & Koellner-Clark, K. (2003). Using a Modeling Approach to Analyze the Ways in Which Teachers Consider New Ways to Teach Mathematics. *Mathematical Thinking & Learning, 5*(2/3), 191-210.

Schorr, R. Y., & Lesh, R. (2003). A modeling approach for providing teacher development. In R. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: A models & modeling perspective on mathematics problem solving, learning and teaching* (pp. 141–157). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Smith III, J., & Thompson, P. W. (2008). Quantitative reasoning and the development of algebraic reasoning. In J. J. Kaput, D. W. Carraher & M. L. Blanton (Eds.), Algebra in the Early Grades (pp. 95-132). New York, NY: Lawrence Erlbaum Associates.

Sriraman, B. & English, L. (2010). Surveying theories and philosophies of mathematics education. In B. Sriraman & L. English (Eds.), *Theories of mathematics education*. Seeking new frontiers (pp. 7–32). Berlin: Springer.

Sztajn, P., Marrongelle, K., & Smith, P. (2011). *Supporting Implementation of the Common Core State Standards for Mathematics: Recommendations for Professional Development*. Retrieved on February 20, 2012 from http://hub.mspnet.org/index.cfm/24233

Thompson, P.W. (1990). A theoretical model of quantity-based reasoning in arithmetic and algebraic. Unpublished manuscript, Center for Research in Mathematics & Science Education, San Diego State University.
Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlin & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education*. WISDOMe Mongraphs (Vol. 1, pp. 33-57). Laramie, WY: University of Wyoming.
Thompson, P. W. (2013). In the absence of meaning… . In Leatham, K. (Ed.), *Vital directions for research in mathematics education* (pp. 57-93). New York, NY: Springer.

AN OBSERVATION INSTRUMENT FOR ASSESSING THE K-16 MATHEMATICS CLASSROOM

Jim Gleason, University of Alabama

We describe the development of a new observation protocol instrument for classroom instruction that is mathematics-specific, spans K-16 mathematics, improves validity and reliability compared to existing instruments, and encompasses the Standards for Mathematical Practice. The instrument may be helpful for educators/researchers engaged in classroom evaluations of K-16 mathematics teaching.

Key words: Classroom research, observation protocol, instrument development

Background

One of the most widely used instruments for mathematics instruction is the Reformed Teaching Observation Protocol (RTOP), developed by the Evaluation Facilitation Group (EFG) of the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT). It was designed to capture "reformed teaching" in mathematics and science. The RTOP was developed and validated based on 141 observations, but only 38 of these (approximately 27%) were mathematics lessons and the remainder were science (Sawada et al, 2002). Furthermore, some of the language in the RTOP reads as science-specific making it unclear as to how to interpret and apply some items to mathematics instruction (e.g., Item 12: "Students made predictions, estimations and/or hypothesis and devised means of testing them."). Finally, the factor loadings indicate that the majority of the variance loaded onto factor one ("inquiry orientation"), with only one item loading uniquely in factor three (Piburn &Sawada, 2000).

Our team, comprised of three mathematics educators, a mathematician, and an educational statistician, addressed these areas to develop a new observation instrument that was: 1) focused on the essential and unique aspects of standards-based mathematics instruction, 2) measured three foundational classroom components underlying the Standards of Mathematical Practice, 3) based on observations of and designed for use in K-16 mathematics classrooms, 4) increased reliability and validity, 5) distributed loadings across three factors.

Theoretical Framework

The Common Core State Standards initiative has one unifying aspect across all K-12 grade bands in the Standards for Mathematical Practice. The Standards of Mathematical Practice define the expectations for all teachers of mathematics to develop mathematical proficiencies in their students (CCSS, 2010). The eight practice standards began with the NCTM process standards of problem solving, reasoning and proof, communication, representation, and connections (NCTM, 2000) and finished with the NRC's proficiency strands of adaptive reasoning, strategic competence, conceptual understanding, procedural fluency, and productive disposition (NRC, 2001). The merging of the NCTM process standards and the NRC proficiency strands have resulted in the Standards of Mathematical Practice and the belief that students should see mathematics as sensible, useful, worthwhile, and that they will believe in their own diligence with a positive self-efficacy. Though these standards usually are in the context of the K-12 classroom, they are easily generalized to the undergraduate mathematics context.

When our team examined the eight practice standards, we found three overarching themes of what a classroom should "look like". These three themes (or constructs) represent (1) how students are engaged during class and with what they are engaged, (2) a well-designed lesson enacted in a manner to promote deep mathematical understanding, and (3) a classroom culture of respect that generates positive productive discourse for students with both teachers and their peers. These three key elements of student engagement, lesson design and implementation, and class culture based in discourse are essential to enacting the practice standards well. While it is difficult to implement all eight practice standards daily, these three themes are crucial for current and future teachers during their continual development in the teaching profession. Therefore, the development of our instrument was grounded in the eight practice standards, which can be found within the final instrument.

Educational/Scientific Importance

There are various ways to categorize events that unfold in the course of teaching a mathematics lesson. The RTOP assesses inquiry orientation, content propositional knowledge, and collaboration; the Mathematical Quality of Instruction (MQI) protocol assesses the relationships between the teacher, student, and mathematical content; and the UTeach Observation Protocol (UTOP) creates an in-depth analysis of the math or science classroom. Additional vantage points, such as those used by this instrument, could provide further insight into demystifying and naming what effective teaching of mathematics entails. This instrument is unique because it assesses the three underlying classroom components of the Standards of Mathematical Practice (student engagement, lesson design and implementation, and class culture based in discourse). To date, the mathematics education community lacks a classroom observation instrument with the vantage point grounded in the underlying constructs of the Standards for Mathematical Practice. This continues to be a timely and pressing need in mathematics education research and teacher education given the goal of successful, large-scale implementation of the Common Core State Standards nationally. As the Science and Mathematics Teacher Imperative (SMTI)/The Leadership Collaborative (TLC) Working Group on Common Core Standards (CCSS) stated, "we simply do not know enough about the effective instructional strategies for teaching…mathematics to all students" (APLU, 2011, p.6). The development of this instrument for use in K-16 mathematics instruction with the aim of generating a mathematics-specific, valid and reliable observation instrument which can now be used to contribute to a) future research on classroom teaching, b) provide feedback to preservice programs, c) improve the transitions within K-16 mathematics, and (d) guide mathematics instruction across K-16 classrooms.

Methods

Instrument Development

Using the RTOP as an initial guide, the research team created and/or revised items to fit the Standards of Mathematical Practice and the three constructs of student engagement, lesson content, and classroom culture and discourse. These items were then revised following use analyzing classroom videos. Additionally, a training manual was developed to improve inter-rater reliability and fidelity to the item intent.

Data Collection

Two types of data were collected in the analysis of the protocol instrument. The first set of data relates to the face validity of the protocol with the items sent out to practitioners in the field of mathematics education to determine if the mathematics education community

believes that the items measure important components of the mathematics classroom and align with the standards of mathematical practice.
The protocol instrument was then used in a variety of undergraduate mathematics classrooms at a large public university in the southeastern United States. These classrooms varied in size from 15 to 150 and in level from College Algebra to Introductory Real Analysis.

Analysis
The analysis of the data includes a factor analysis to determine the construct validity of the protocol instrument and a regression analysis using a subset of the observed classrooms to determine the predictive validity of the protocol instrument on student achievement.

Questions for Audience:

- What are some uses for such an observation protocol for research in undergraduate mathematics education?
- Are there essential components of the mathematics classroom that this protocol is missing?
- Are there additional data points needed or additional analyses that should be performed on the data?

References

Common Core State Standards (CCSS). (2010). *The Common Core State Standards Initiative: Preparing America's students for college and career.* The Council of Chief State School Officers and the National Governors Association Center.
National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics.* Reston, VA: Author.
National Research Council. (2001). *Adding it up: Helphing children learn mathematics.* Washington, DC: National Academy Press.
Piburn, M., & Sawada, D. (2000). *Reformed Teaching Observation Protocol (RTOP): Reference Manual*. Arizona State University. Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers. Retrieved from http://www.public.asu.edu/~anton1/AssessArticles/Assessments/Chemistry%20Assessments/RTOP%20Reference%20Manual.pdf
Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R. & Bloom, I. (2002), Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics*, 102: 245–253. doi: 10.1111/j.1949-8594.2002.tb17883.x
Association of Public and Land-grant Universities (2011). *The common core state standards and teacher preparation: The role of higher education*. APLU/SMTI, Paper 2. Washington, DC: Association of Public and Land-grant Universities. http://www.aplu.org/document.doc?id=3482

An Investigation of College Students' Statistical Literacy

Erin Glover
Portland State University

Sean Larsen
Portland State University

Introduction

Statistics educators consider statistical literacy a vital skill because it supports students in thinking critically about the way data is used in everyday social, political and medical contexts. Statistical literacy goes beyond simply reading graphs to include interpreting their meaning and evaluating, with a questioning attitude, the information being presented (Shaughnessy, 2007; Gal, 2002; Watson & Moritz, 1997). Watson (1997) describes statistical literacy in terms of a three-tiered hierarchy delineating the skills necessary for interpreting stochastic information.

Tier 1: a basic understanding of probabilistic and statistical terminology
Tier 2: an understanding of probabilistic and statistical language and concepts when they are embedded in the context of wider social discussion
Tier 3: a questioning attitude, which can apply more sophisticated concepts to contradict claims made without proper statistical foundation\

The researchers conducted a classroom teaching experiment in an introductory college statistics course that aimed to support the development of the skills described in Watson's model. This research report investigates advances in students' statistical literacy, focusing both on advances in their language and usage of statistical concepts, and advances in their ability to contextualize situations that would produce a given data set. In the teaching experiment, students were given a set of graphs and asked to interpret them, compare and contrast them, and to describe real life contexts that might explain differences between them. Four weeks later in the course students revisited these same graphs. Our analysis of student responses highlights students' development in statistical literacy, in particular, their advances in using statistical language correctly in real-world contexts.

Theoretical perspective

Watson's model describes three categories of skills required to fully understand, interpret, and communicate stochastic information. The instructional approach was designed to develop the skills described in all three categories. While students were learning statistical terminology (Tier 1) they were simultaneously engaged in exploratory data analysis (Tier 1 and Tier 2), creating and critically interpreting statistical models (Tier 1 2, and 3), using their newly acquired terminology to describe and support their findings. These statistical literacy tasks required students to engage in sophisticated reasoning about distributions of data. The analysis of students' reasoning was guided by the work of Noll & Shaughnessy (2012).

Noll & Shaughnessy's (2012) presented a framework meant to support characterizing students' reasoning about sampling concepts. The hierarchical framework characterized student reasoning primarily in terms of three levels. Additive reasoning described frequency-based reasoning in student responses (e.g., "there are more red ones"). The second level includes four categories of single-attribute reasoning: Weak Center, Shape, Variation, and Proportional (Strong Center). Weak Center reasoning is indicated by attention primarily on modes while Strong Center reasoning is indicated by responses referred to population proportions (percentages, means, or medians). Distributional reasoning involves the coordination of two or more of the (second level) single-attribute characteristics.

Methods

Data Collection: The author implemented the CATALST curriculum materials (Garfield, delMas, & Zieffler, 2010) using TinkerPlots^{TM} software in an undergraduate introductory statistics course at a large commuter university. Students enrolled in this course as a prerequisite for the traditional statistics sequence, or to satisfy the required math elective needed to graduate. A total of 21 students enrolled in the course and all students consented to be participants in the study. The data collected consisted of student work and video recorded engagement with tasks during in-class activities and individual interviews. This analysis will focus on work from a homework activity completed on week 2 and an exam taken during week 6 which included a follow-up to the week 2 activity.

Data Analysis: The process of characterizing student responses in the teaching experiment drew on the Noll & Shaughnessy framework (2007). While this framework emphasizes the distinction between Additive and Proportional reasoning, the focus here was on characterizing student responses according to how they analyzed and contextualized graphical information. The analysis of student work uncovered three main characteristics addressed in student responses, which were coded as Variation, Shape, and Location/Placement. When students were able to coordinate more than one of these characteristics, a Distributional Reasoning code was used. The Variation, Shape, and Distribution codes were consistent with the corresponding categories of the Noll & Shaughnessy framework, while the Location/Placement code emerged from the analysis in order to capture students' attention to *where* the data was located in a distribution.

Application of the Variation, Shape, and Distribution codes in the analysis of this work was consistent with the Noll & Shaughnessy's definitions of these constructs. The framework characterized student responses as *Variation* if they included concepts such as range, standard deviation, interquartile range and variability. Additionally, in this study the use of words such as, "spread" (from the center), "cluster", or "outlier" resulted in a variation code. The Noll & Shaughnessy framework characterized student responses with *Shape* if language such as "skewness", "normally distributed", "bell curve", "evenly distributed", "smooth", or "bumpy" was used to describe distributions. Seeing similar responses in the analysis of the work presented in this paper, student arguments based on these characteristics were coded as *Shape*. For the purpose of this research, both statistical terminology (like "normal distribution") and naïve language (like "gaps" or "smooth") was also included in the Shape code. The Noll & Shaughnessy framework identified student responses that coordinated one or more of the attributes from Shape, Variation, or Proportional reasoning were described to be "Distributional" reasoning. Similarly, student responses that coordinated two or more characteristics within the Shape, Variation or Location/Placement codes were coded as "Distributional Reasoning" in this analysis.

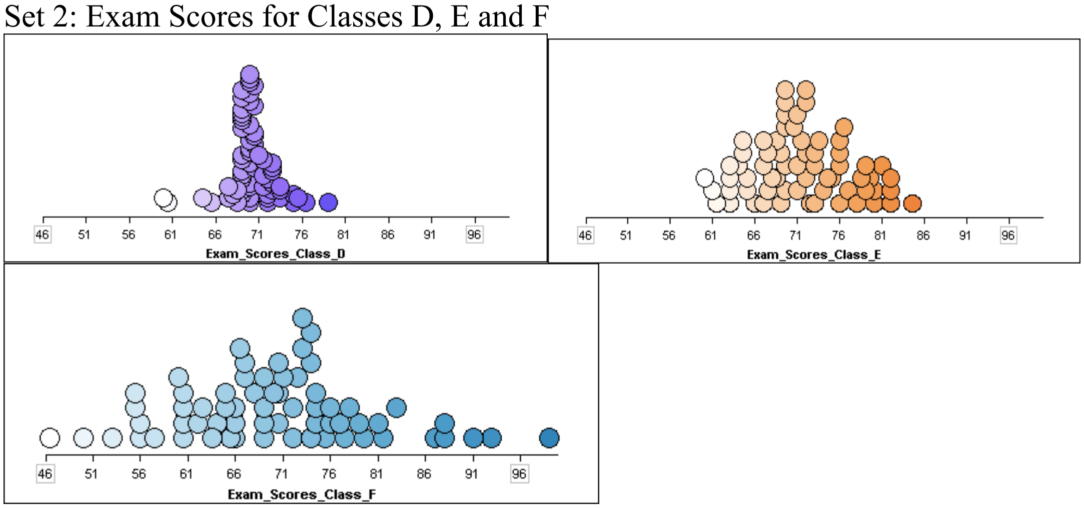
The curriculum often prompted students do contextualize information presented in graphical information, so it became necessary to code these responses separately into two categories: Reasonable or Unreasonable. If the student did not provide an appropriate context (Unreasonable) a Level-0 code was given. If the context provided by the student could be coded as reasonable, then one of three sub-codes was used to describe their response. If the response gave a general context that fit the situation but did not speak to the graph characteristics, a Level-1A code was given. A level-1B code was given if the student responded to context-based problems with descriptions of statistical measures. If the student met the criteria for both Level-1A and Level-1B codes, then a Level-2 code was used.

Results

The results will be presented in two parts. The first will be the analysis of student responses around distributional thinking. The second will be the student results around contextualizing graphical information. The analysis will show marked improvement in both categories, but will be treated separately.

Distributional Thinking. In the second week of the course students were presented with two sets of three graphs in a homework activity that asked the students to compare and contrast graphs and provide real-life contexts that might explain differences between them. The students were told that the graphs represented exam scores for different classes and provided no other information about them. In the sixth week of the course, students revisited these same graphs during their midterm exam. They were asked again to compare and contrast the graphs, but they were also asked to state which class they would prefer to be in and why.

Set 2: Exam Scores for Classes D, E and F



In the second week of the course, students' comparisons were often focused on the location and/or placement of the data. Many of the students listed statistical measures, but lacked coordination of the statistical concepts with their contextual statements. By the sixth week of the course, students' responses coordinated the location and placement of data with variation and/or shape arguments as well.

The categories of student reasoning of interest here can be illustrated by considering the responses given by one student, Aaron. In Week 2, Aaron described the characteristic that distinguished the three graphs to be "the mode of each classes' exam scores." This response was coded with a Location/Placement code. Revisiting the same graphs in Week 6, Aaron's comparisons used more sophisticated language and included statistical concepts.

Aaron: Class D has the highest variation, but shares about the same variability with Class E. That is, in Class D, while the data fall off from the typical value quicker than in any other class, the range of variability is almost the same with Class E. The range in Class D is 20, while Class E's is 25. This is notable only because Class F's range of variability is much greater, sitting at about 50. The centers of all the graphs are almost the same, at around 71. The shapes of all three either condense around 71 or fall out from it.

Aaron's Week 6 response was coded with Location/Placement, Variation and Shape. Because Aaron was coordinated these three kinds of characteristics, his response was also coded as Distributional reasoning.

Examining classroom results for the second set of graphs, eight of 21 students coordinated at least two of Shape, Location/Placement and Variation in their Week 2 responses resulting in a Distributional Reasoning categorization. This number increased to 17 by Week 6.

Week 2 Code Counts		**Week 6 Code Counts**	
Shape=4 Variation=14 Location/Placement=11	Distributional Reasoning= 8 Other =2	Shape=10 Variation=17 Location/Placement=18	Distributional Reasoning=17 Other =2

Context Ideas. To illustrate our analysis of students' contextualizing of graphical information, we look again to Aaron's responses to the 2nd set of graphs. In Week 2, Aaron provided a Level-1A response since he provided an appropriate context, but did not use statistical measures in his argument.

Adam: *Perhaps in class D, where most of the students received the same score, more of the students studied together for the exam, and therefore wrote down the same answers on their tests.*

Looking forward, during week 6 Adam received a level-2 code because he uses reasonable contextual language and ideas (Level-1A) for a particular graph, but also uses statistical measures (Level-1B) to form his argument.

Adam: If we're talking about a curved exam here, I'd be best off in Class F where everyone is getting nearly the same grade. Because people are not likely to get a grade significantly above a 71, I'd be best with a score of 71.

Examining classroom results for the 2nd set of graphs, there was notable improvement in students' ability to contextualize graphical information. In Week 2, nine of the 21 students used Level-2 context responses. That number increased to 17 in Week 6.

Week 2 Code Counts		Week 6 Code Counts	
No Response= 3 Level-0=1	Level-1= 8 Level-2= 9	No Response= 3 Level-0=0	Level-1= 1 Level-2=17

Conclusions and Directions for Future Research

It is clear that the students exhibited improvement in their responses when asked to compare the same sets of graphs in week 2 and again during week 6 of the course.

Students used more sophisticated reasoning later in the course, and also showed progress in contextualizing graphical information. While this research does not make claims about what exactly led to students' improvement, the analysis contributes to the statistics education literature in three key ways. First, the analysis of student work produced a category of student reasoning which compliments the Noll & Shaughnessy framework. Students' consideration of centers are categorized by the Noll & Shaughnessy framework as either "weak center" or "proportional" (strong center) depending on whether they involve proportional reasoning (mean) or not (mode). Our analysis did not focus on this distinction, but rather revealed that students used a variety of notions including centers to describe "where" a distribution was positioned. We observed students providing sophisticated explanations involved the coordination of such notions of location/placement and other single-attribute characteristics such as shape and variation. Secondly, the tasks of comparing graphs and providing contexts to support the graphical information seem particularly effective at eliciting expressions of students' reasoning. Notice how Aaron's responses give us detailed information about how he reasoned about the graphs contextually using distributional thinking. This suggests that activities like this can be useful for researchers interested in studying statistical literacy. Finally, because the tasks elicited such sophisticated student reasoning, it can be conjectured that repeated engagement in similar tasks could support development of statistics literacy. Future analyses of the data set will be focused on investigating whether and how the curriculum and instructional approach supported students' learning.

References

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70,* 1–25.

Garfield, J., delMas, R, & Zieffler, A. (2010). Developing Tertiary-Level Students' Statistical Thinking through the Use of Model-Eliciting Activities. *Proceedings of the 8th International Conference on Statistics Education* (ICOTS-8), Lubljana, Slovenia.

Noll, J., & Shaughnessy, J. M. (2012). Aspects of students' reasoning about variation in empirical sampling distributions. *Journal for Research in Mathematics Education*, *43*(5), 509-56.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. *Second Handbook of Research on Mathematics Teaching and Learning*, 957-1009.

Watson, J. (1997). Assessing statistical thinking using the media In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107-121).

Watson, J. M., & Moritz, J. B. (1997a). Student analysis of variables in a media context. In B. Phillips (Ed.), *Papers on Statistical Education Presented at ICME 8* (pp.129–147). Haw- thorn, Australia: Swinburne Press.

Characteristics of Successful Programs in College Calculus: Instructors' Perceptions of the Usefulness and Role of Instructional Technology

Erin Glover
Portland State University

Sean Larsen
Portland State University

Abstract

The CSPCC (Characteristics of Successful Programs in College Calculus) project is a large empirical study, investigating mainstream Calculus 1, that aims to identify the factors that contribute to successful programs. The CSPCC project consists of two phases. Phase 1 entailed large-scale surveys of a stratified random sample of college Calculus 1 classes across the United States. Phase 2 involves explanatory case study research into programs that were identified as successful based in part on the results of the Phase 1 survey. This second phase will lead to the development of a theoretical framework for understanding how to build a successful program in calculus and in illustrative case studies for widespread dissemination. Technology was one of the topics we explored with students, instructors, administrators, and other individuals that we interviewed during our case study site visits. In this preliminary report, we will focus on calculus instructors' views on instructional technology.

Study Background and Research Question

The CSPCC (Characteristics of Successful Programs in College Calculus) project is a large empirical study, investigating mainstream Calculus 1, that aims to identify the factors that contribute to successful programs. The CSPCC project consists of two phases. Phase 1 entailed large-scale surveys of a stratified random sample of college Calculus 1 classes across the United States. Phase 2 involves explanatory case study research into programs that were identified as successful based in part on the results of the Phase 1 survey. Specifically, institutions were selected based on student *persistence* (continuing on to take Calculus 2), *success* (pass rates in Calculus 1), and reported increases in students' *interest*, *confidence*, and *enjoyment* of mathematics as a result of taking Calculus 1. This second phase will lead to the development of a theoretical framework for understanding how to build a successful program in calculus and in illustrative case studies for widespread dissemination.

Technology was one of the topics we explored with students, instructors, administrators, and other individuals that we interviewed during our case study site visits. The questions we asked were designed to explore the role of technology in successful calculus programs. ***In this preliminary report, we will focus on calculus instructors' views on instructional technology***.

Relation of this work to the research literature

Incorporating technology into mathematics education has been of interest for the last few decades. During the calculus reform movement in the 90's, mathematics researchers turned their attention towards how technology could be used effectively in teaching, and what approaches would be successful. An abundance of new technologies have become available for teachers and students alike, so there is continued interest how technology can be leveraged to improve mathematics instruction (Dick, 2007). Researchers have noted that computational technology (calculators and CAS) can allow students to focus on conceptual knowledge rather than procedural skills (Demana, 1990). More recently, researchers have been looking at how to integrate more dynamical technology (e.g., GeoGebra) to support students' visualization (Simonsen & Dick, 2007). As Simonsen & Dick note, "The use of technology in mathematics classrooms raises several areas of concern for teachers: curriculum issues, classroom dynamics, training and support, and technological accessibility." It stands to reason that teacher's perceptions of technology may have a major impact on how the potential of these tools can be realized. Our analysis will contribute to the research literature around educational technology by providing characterizations of instructors' perspectives about technology in institutions identified as having successful Calculus 1 programs.

Research Methods and Analytic Framework

Our research team conducted site visits at five bachelors granting institutions. Four of the institutions were universities (two had very recently transitioned from colleges to universities). Three of these universities were private, while the fourth was a large urban public university. The final case study institution was a private liberal arts college. While on campus, we interviewed students, instructors, administrators, and others involved in the calculus program at the institution. This report will focus on the instructor interviews. We interviewed a total of 25 instructors over the course of the five case study site visits. Here, we will report on our ongoing analyses of the instructors' views about technology as revealed by both the instructors' responses to questions directly about technology and spontaneous comments about technology in response to other questions.

The first stage of our data analysis involved the larger CSPCC project team and the full collection of interviews conducted with students, instructors, department chairs, calculus coordinators, and selected administrators at all four institution types. This stage involved "tagging" all transcripts by identifying the relevant topics addressed by the interviewee in each of their responses. For example if an instructor said, "some of my students are not prepared to handle the algebraic procedures needed to use the derivative concept on the application problems that we put on our common exams" this response would be tagged with the following codes: *student subject characteristic*, *assignments and assessments*, *course coordination*, and *content*. The CSPCC tagging scheme consists of 24 different tags including: *teaching and learning*, *course structure*, *instructional materials*, and (most relevant for this report) *technology*. This tagging was done using Hyperresearch, which allowed us to run reports of all instances in which technology was mentioned by an interviewee. Two coders tagged each transcript independently and a final version of the coding was created (for each response) by taking the union of the tags used by the two coders. This approach allows us to be confident that a very high percentage of the relevant interview excerpts related to technology have been identified so they can be analyzed in depth.

The second stage of analysis will involve (primarily) open coding (Corbin & Strauss, 2008) of the instructors' statements about technology. However, this process will also be informed by the categories of issues identified by Simonsen & Dick (2007) in their investigation of high school teachers' perceptions of the impact of graphing calculators. Table 1 presents a subset of the major trends that emerged from Simonsen & Dick's analyses of the teachers' perceptions. Our analysis will attend to these themes as well as others that emerge from our data.

Advantages	**Disadvantages**	**Classroom Dynamics**
Less distraction w/ computation	Logistical difficulties	Less teacher-centered
Immediate feedback	Fear of calculator dependency	More open-ended questions
Enhancement of visualization	Time spent learning calculator	More cooperative learning

Results of the research

Survey Results. There were a few questions on the instructor surveys that give us some insight into how much technology use was happening at the selected case study schools. One of the findings from the survey analysis is that instructors at the selected institutions reported significantly greater use of graphing greater use of graphing calculators. In particular, the survey results showed that instructors at selected schools permitted (and required) use of graphing calculators at a significantly higher percentage than the schools that were not selected. Additionally, these instructors reported more demonstrations in class using graphing calculators and more use of graphic calculators by students during class (both of these differences were statistically significant). While we cannot make claims that the use of technology was a cause of success (defined by our selection criteria) at these institutions, these results do present us with a correlation worth investigating further. Additionally, they suggest that, when we analyze our interviews, we will be investigating the views of calculus instructors at successful intuitions where technology use is more prevalent than average.

Sample of Interview Results **and** ***Directions for Ongoing Analysis.*** Simonsen & Dick's (2007) research suggested that technology was used by many instructors to aid in students' visualization of mathematical concepts and providing immediate feedback to students. Preliminary analyses of our interviews suggest that the instructors at the selected case study institutions used technology in and out of the classroom for similar reasons.

"I think it really helps aid their visualization of a variety of things. It's not always relevant, so I don't use it for everything, but often it will draw a graph of this function a lot nicer than I would on the board, and then I can also manipulate that in certain ways."

"…online homework where they get immediate feedback and they can keep doing it many times until it's correct with a few exceptions."

Although technology is often used by instructors to assist students in learning calculus concepts, not all instructors agree that the use of technology, such as calculators or web-based calculators (like Wolfram Alpha) are appropriate for learning calculus concepts. As Simonsen and Dick found in their research on calculators, a prevalent issue was logistical difficulties with implementing the tool. Our preliminary analysis suggests instructors at successful institutions had similar difficulties.

"I felt like it put up an unnecessary wall between my students and the learning because at first there were tons of errors with WeBWorK and they may understand the material, but it like -- enter it in the computer and it just -- they just were angry all the time."

One instructor, in particular, believed that some computational technology did not do students "any service because it gives them a false sense that they know what they're doing."

These preliminary results suggest that some of the perceptions of the teachers studied by Simonsen and Dick (2007) are shared by the instructors at our case study institutions. Further analyses will be dedicated drawing more complete comparisons between these two groups and investigating other issues that may be unique to our population (e.g., instructor driven teaching innovations using technology).

Discussion Questions:

- How could we use our student focus group interviews to address our research question?
- What topics should we search interviews for in conjunction with technology in order to address our research question?
- What kind of analysis will help us to identify potential connections between teachers' views and their teaching practice (and ultimately their students' success)?

References

Corbin, J., & Strauss, A. (Eds.). (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage.

Demana, F. & Waits, B. K. (1990a). The role of technology in teaching mathematics. The Mathematics Teacher, 83(1), 27-31.

Simonsen, L. M., & Dick, T. P. (1997). Teachers' perceptions of the impact of graphing calculators in the mathematics classroom. *Journal of Computers in Mathematics and Science Teaching*, *16*(2), 239-368.

Tall, D., Smith, D., & Piez, C. (2008). Technology and calculus. *Research on technology and the teaching and learning of mathematics*, *1*, 207-258.

Zbiek, R. M., Heid, M. K., Blume, G. W., & Dick, T. P. (2007). Research on technology in mathematics education: A perspective of constructs. *Second handbook of research on mathematics teaching and learning*, *2*, 1169-1207.

PROFESSIONAL DEVELOPMENT AND STUDENT ACHIEVEMENT ON STANDARDIZED STATE EXAMS

Melissa L. Goss	Robert A. Powers	Rebecca A. Dibbs
University of Northern Colorado	University of Northern Colorado	University of Northern Colorado

Although teacher quality is positively correlated with student achievement, easily quantified measures of teacher quality are not accurate measures of quality; teacher pedagogical content knowledge and skills are better predictors, but difficult to measure. Professional development may be a cost-effective vehicle for developing new skills in in-service teachers, but there is conflicting research on whether professional development measurably raises student achievement on high stakes standardized tests. The purpose of this causal-comparative study was to examine Andrew, an in-service, high school teacher participant in the master's program. State mathematics assessment and student demographic data were collected from school districts for 4 academic years spanning from pre-program through program completion. One-way ANOVA analysis on student scale scores factoring by year showed a significant decrease in student mathematics scale scores potentially attributable to differences in population. Independent-samples t tests on the final two years showed a statistically insignificant increase in student growth percentiles.

Key words: in-service teachers, professional development, secondary education, student achievement

The general consensus is that high quality teachers improve student achievement, but demonstrating this association quantitatively as any easily measured credential is not necessarily an accurate measure of teacher quality (Foster, Toma, & Troske, 2013; Rockoff, 2004), and the pedagogical content knowledge, which appears to account for most of the variance in teacher quality, is difficult to measure (Dash et al., 2012). Although professional development can help in-service teachers to examine and improve their practice, the conflicting research on the relationship between professional development and student achievement indicates that this relationship may be dependent on the specifics of a given professional development program (Blank & de las Alas, 2009; Huffman, Thomas & Laurenz, 2003; Ross, Hogaboam-Grey, & Bruce, 2006). Furthermore, even when there is significant achievement gains attributable to professional development, increased student achievement occurs after the professional development has ended (Harris & Sass, 2007).

This study investigates the relationship between professional development and student achievement of participating teachers in a 2-year, blended face-to-face and online delivered master's program in mathematics for in-service secondary teachers. The question guiding this research was: do Math TLC master's program teacher participants' students' state mathematics scores differ between pre-program and post-program controlling for demographic variables and teachers' pedagogical content knowledge?

Methods

Setting/Participants

The Math TLCmaster's program is a 2-year, blended face-to-face and online delivered program for in-service secondary teachers. Offered through a joint effort at two Rocky Mountain region universities, cohorts of 16 to 20 new teacher participants each year complete a 2-year master's program in mathematics with an emphasis in teaching (about half of course credits in mathematics, half in mathematics education). The primary goals of the program are to develop content proficiency, cultural competence, and pedagogical expertise for the teaching of secondary mathematics. So far, 31 teachers from 3 cohorts have successfully completed the program.

This is a quantitative causal-comparative study (Gall, Gall, and Borg, 2007) on the state scores of the students of one teacher participant of the master's program. We have collected the participant's students' demographic and state assessment data for four years. We also have conducted quantitative observations of the teacher participant's teaching both before and after his enrollment in the program (Hauk, Jackson, & Noblet, 2010; Goss, Powers, & Hauk, 2013) measures of his pedagogical content knowledge for teaching using a written instrument (Hauk, Toney, Jackson, Nair, & Tsay, 2013) throughout his participation in the program, and measures of his intercultural competence pre-program and post-program. We focus this report on the student outcomes from one teacher participant from the first cohort.

The participant for this study is Andrew who began the program in the summer of 2009 and completed it in 2011. Andrew started the program teaching at a high school of approximately 1500 students within a district of 19,000 students with a high school graduation rate of 72.4%. Two years into the program, Andrew transferred to another high school of 1100 students within a new district of 4600 students with a graduation rate of 86%. Andrew taught grades 9-12, but only 9th and 10th graders complete the state assessment.

We considered the state assessment scores of Andrew's 9th and 10th grade students from four academic years: (0) prior to his beginning the program, (1) his first year in the program, (2) his second year in the program, and (3) the year following his completion of the program. Prior to his beginning the program, Andrew taught 15 students in 9th grade, 94 students in 10th grade in the course of Geometry which typically is taught at the 10th grade level indicating his 9th graders this year were in an advanced class. In his first year of the program, Andrew taught 54 students in 9th grade and 76 students in 10th grade in the courses of Algebra 1 and Geometry. In his second year in the program, he taught 35 students in 9th grade and 21 students in 10th grade in the courses of Algebra 1 and Algebra 2. In the year following his completion of the program, he taught 62 students in 9th grade and 5 students in 10th grade in the courses of Algebra 1 and Everyday Math.

The Colorado state assessment is administered to students in grades 3-10 in the subjects of mathematics, English/Language Arts, and science in the spring of each academic year. For each subject, students receive a scale score, a proficiency level, and growth percentile rating. Scale scores are conversions from raw scores that represent the same level of achievement regardless of the year in which the test was administered (Colorado Student Assessment Program, 2011). The growth percentile reports each student's progress by comparing each student's current achievement to students in the same grade throughout the state who had similar scores in past years and is only reported if the student completed the assessment in two consecutive years.

The writers of the state assessment strived to establish content validity by having content-area specialists, teachers, and assessment experts develop a pool of items that measured the state's assessment frameworks in each grade and content area. The state reports the 9th and 10th grade mathematics test showed good internal consistency (Cronbach's alphas range from .91 to .94) and interrater reliability on constructed response items showed kappa values ranging from .60 to .93 (Colorado Department of Education, 2009, 2010, 2011, 2012).

Data Collection

From the participants' districts we requested both teacher and student data. We requested students' demographic data, state scale mathematics scores, state mathematics proficiency levels, student growth percentiles, and state English and Language Arts scale scores. For each teacher, for each of the given years, we requested all data for students in that year plus all data for those same students for the previous year. We have obtained students' mathematics scale scores from Andrew's first two years; these scores have not been linked to students' previous year's scores or their growth percentiles. We have obtained all requested variables from the second two years.

Data Analysis

Andrew's students' scale mathematics scale scores for the four years requested have been collected. We conducted one-way ANOVA analyses on the mathematics scale scores; the students enrolled in Andrew's classes in the four academic years composed the four groups. We conducted post-hoc Tukey analyses. For the two years that student growth percentiles have been collected, we conducted independent-samples *t* tests on those growth percentiles. Note that these growth scores were not available for all students enrolled.

Results

Table 1 presents the mean and standard deviation for each year for Andrew's students' state scores. There was a decrease in scores between the year prior to entering the program and the first year of his program at his first school. After transferring to his second school, there was also a decrease in scores.

Table 1

Descriptive Statistics Student Mathematics Scale Scores

Year	*N*	*M*	*SD*
Prior to entering program (0)	109	610.03	66.83
First year in program (1)	130	582.83	73.52
Second year in program (2)	65	592.02	58.72
After completing program (3)	68	550.87	52.41

An ANOVA shows a significant difference in Andrew's students' scale mathematics scores across his years of enrollment in the master's program (see Table 2). Post-hoc Tukey tests presented in Table 3 confirm that the decreases in mean scale mathematics scores are statistically significant.

Table 2

Summary of ANOVA

	SS	*df*	*MS*	*F*	*p*
Between Groups	148838.25	3	49612.75	11.473	<.001
Within Groups	1543779.97	357	4324.31		
Total	1692618.22	360			

Table 3

Tukey HSD Comparison for Number of Years Since Beginning the Master's Program

(I) Years	(J) Years	Mean Difference (I-J)	*SE*	*P*
0	1	27.20	8.54	.009
	2	18.01	10.88	.349
	3	59.16	10.21	.000
1	2	-9.19	10.58	.821
	3	31.87	11.97	.007
2	3	41.15	11.97	.004

We also conducted independent-samples *t*-tests on the growth percentiles for Andrew's students from his two years in the second district. We found an increase in growth percentile between year 2, his last year in the program (M = 46.35, SD = 25.52) and year 3, the year following his completion of the program (*M* = 50.18, *SD* = 30.21), but it is not statistically significant: $t(112) = 0.72$, $p = .471$.

Discussion

Andrew's students' scale mathematics scores significantly decreased across years. This potentially can be attributed to the difference in student populations across years which cannot yet be accounted for due to the unavailability of students' previous years' scores. In the first year analyzed, Andrew taught geometry classes that included a group of 9th grade students advanced beyond their grade level. In the year following his completion of the program, he taught Algebra 1 and a class of Everyday Math, a course typically designed as an intervention for students below grade level. Alternatively, the literature suggests a significant drop in student scores can be expected as significant rises in scores are typically not seen until two years after the completion of professional development (Blank & de las Alas, 2009).

Although statistically insignificant, the increase in growth percentiles across the two years analyzed indicates more research on the growth percentile variable across all four years is warranted. Continuing data collection efforts strive for students' current years' test scores to be linked with their scores from previous years as well as collecting state calculated growth percentiles.

We anticipate the collection of student data for 20 of the teacher participants who have completed the program. We plan to analyze student assessment scores controlling for previous test scores and other student demographic and teacher data. Past research team efforts have collected scores for pedagogical content knowledge and intercultural competence, we plan to use linear modeling to see how these variables, time in the program, and student demographic variables contribute to student state mathematics assessment scores.

1. Previous research team efforts have collected measures of teacher participant PCK and intercultural competence. What other statistical models should we try?
2. Requesting and collecting data from K-12 school districts is difficult as the data professionals employed by the district do not often have time built into their schedules for such data-pulls. How can we better facilitate data-sharing?
3. Past literature has indicated that there is often a lag between completion of PD and implementation of what was learned into the classroom. Should we continue to collect data on our same teachers in years to come?

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. DUE0832026. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Blank, R. K., & de las Alas, N. (2009). Effects of teacher professional development on gains in student achievement. *Peabody Journal of Education*, *77*(4), 59-85.

Colorado Department of Education (2009, 2010, 2011, 2012). Colorado State Assessment Program Technical Report. Available at http://cde.co.us.

Colorado Department of Education (2012). Transitional Colorado Assessment Program Technical Report. Available at http://cde.co.us.

Dash, S., de Kramer, R. M., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of Online Professional Development on Teacher Quality and Student Achievement in Fifth Grade Mathematics. *Journal of Research on Technology in Education*, *45*(1), 1-26.

Foster, J. M., Toma, E. F., & Troske, S. P. (2013). Does Teacher Professional Development Improve Math and Science Outcomes and Is It Cost Effective?. *Journal of Education Finance*, *38*(3), 255-275.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007) Educational research: An introduction. Boston: Pearson Education.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, *95*(7), 798-812.

Hauk, S., Jackson, B., & Noblet, K. (2010). No teacher left behind: Assessment of secondary mathematics teachers' pedagogical content knowledge. *Proceedings for the 13th Conference on Research in Undergraduate Mathematics Education (electronic).* PDF available at sigmaa.maa.org/rume/crume2010/Archive/HaukNTLB2010_LONG.pdf

Hauk, S., Toney, A., Jackson, B., & Nair, R., Tsay, J.-J. (2013). Illustrating a theory of pedagogical content knowledge for secondary and post-secondary mathematics instruction. *Proceedings for the 16th Conference on Research in Undergraduate Mathematics Education (electronic).*

Huffman, D., Thomas, K., & Lawrenz, F. (2003). Relationship between professional development, teachers' instructional practices, and the achievement of students in science and mathematics. *School Science and Mathematics*, *103*(8), 378-387.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247-252.

Ross, J. A., Hogaboam-Gray, A., & Bruce, C. (2006). The impact of a professional development program on student achievement in grade 6 mathematics. *Journal of mathematics teacher education*, *9*(6), 1-27.

67

DO EXPERTS AND NOVICES GESTURE DIFFERENTLY?

Brent Hancock, Marki Dittman, & Hortensia Soto-Johnson
University of Northern Colorado

Previous gesture studies conjecture that as individuals develop expertise in a field of mathematics their gestures tend to become more metaphoric, iconic, and dynamic. In this mixed-methods study, we compared the gestures of six experts and four pairs of novices as they geometrically described the complex number arithmetic operations z+w, zw, and 1/z. An ANOVA revealed that the factors Task and Gesture were statistically significant, but there was no statistically significant difference between the two groups' gesture use. A Hierarchical Cluster Analysis directed the qualitative analysis where we found that novices exposed to technology appeared to produce gestures that were innovative or similar to the experts' gestures. These findings suggest that facilitating students' awareness of their own and the instructors' gestures as well as exposing students to technology may help them develop more dynamic gestures and in turn possibly facilitate a more geometric perspective of the arithmetic of complex numbers.

Keywords Complex variables, Embodied cognition, Expert, Gesture, Novice

Introduction

The purpose of this mixed methods study is to explore two areas of research: the study of gesture and the field of complex variables. Gesture has become an increasingly popular topic in the field of mathematics education, specifically as a tool to interpret students' understanding or misunderstanding of a mathematical notion. Goldin-Meadow (2003) summarized much of the literature on gesture; however, these studies tend to center on elementary students as they work on arithmetic tasks. In fact, few researchers have explored the role of gesture in undergraduate-level mathematics. Exceptions include Marghetis and Núñez (2010), who explored graduate students' use of gesture as they proved a fixed point theorem, and Sinclair and Tabaghi (2010), who investigated mathematicians' use of gesture while explaining their conception of eigenvectors. Our research differs from these studies in that we compared experts' and novices' use of gesture as they worked on tasks related to arithmetic operations of complex numbers. Our research questions were:

1. Is there a statistically significant difference in the number and types of gestures exhibited by experts and novices as they explained their geometric perception of complex number arithmetic?
2. What is the nature of experts' and novices' use of gesture as they explain their geometric perception of complex number arithmetic?

Review of Literature

Despite the belief that gesture may provide insight into mathematical cognition, the variety of research on the subject is limited. A large portion of mathematics education research articles related to gesture center on mathematical tasks commonly encountered in the primary grades (Alibali, Bassok, Solomon, Syc & Goldin-Meadow, 1999; Goldin-Meadow, Kim & Singer, 1999; Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001; Perry, Church, Goldin-Meadow, 1988). However, some researchers have investigated individuals' conceptions of more advanced mathematical topics in ways that are pertinent to our research. For instance, Marghetis and Núñez (2010) demonstrated how mathematical constructs whose definitions suggest a static object are in fact perceived and communicated through individuals' gestures as dynamic entities. Additionally, Sinclair and Tabaghi (2010) demonstrated the kinetic and temporal characteristics of experts' conceptions of eigenvectors. Both of these studies suggest that a dynamic conception of mathematical constructs appears to be an inherent characteristic of *reification* (Sfard, 1992) of a mathematical notion.

While the overlap of literature on gesture and advanced mathematics is minimal, even less research (Danenhower, 2006; Panaoura, Elia, Gagatis & Giatilis, 2006; Nemirovsky, Rasmussen, Sweeney & Wawro, 2012) exists regarding understanding of complex numbers. Panoura et al. investigated high school students' ability to solve complex arithmetic equations or inequalities using either a primarily algebraic or primarily geometric approach. Panoura et al. found that students who used a geometric approach were more successful at correctly completing tasks than those who used an algebraic approach. In a similar study, Danenhower investigated undergraduate mathematics majors' willingness and ability to switch between different forms used for expressing complex numbers. He characterized students' level of understanding based on their ability to use a single form, represent an expression in different forms, translate between forms, and judge when to shift from one form to another. One pertinent result from this study was that the students held an *object* understanding of the algebraic and vector forms, but only a *process* understanding of the exponential form. Both of these studies indicate students' inflexibility or inability to navigate between representations and forms of complex numbers and equations involving complex numbers, a finding that our study substantiates.

Nemirovsky et al. (2012) investigated preservice teachers' conceptions of complex arithmetic, specifically their geometric interpretation of the addition and multiplication of complex numbers. In the study, students used tape, string, and stick-on dots, in conjunction with a tiled floor in order to invent ways to perform complex addition and multiplication tasks. Students in the study expanded their own "realm of possibilities," (p. 291) that is, the collection of all possible outcomes associated with some perceptuo-motor activity for complex arithmetic as they utilized their environment to enact specific operations, such as multiplication by *i*. This work offers promising teaching techniques for developing an *object* view of the addition and multiplication of complex numbers and informed our choice of a theoretical perspective.

Theoretical Perspective

It is important for researchers investigating gesture to adopt a theoretical perspective that addresses the multifarious contributions from language, social interaction, and the embodiment of human activities and thoughts, all of which shape individuals' gestures. McNeill (2005) argued that gestures fuel thought and speech and that "gestures, language and thought are all different sides of a single mental/brain/action process" (p.1). This Vygotskian perspective necessitates viewing gestures as part of language itself. Given social constructivism's emphasis on language and discourse

(Miller, 2009), McNeill's perspective is appropriate for investigating gesture because it claims that gestures are inseparable from language. However, while social constructivism illuminates many desirable parallels between communication and thought, it does not account for the belief that gestures are both bodily and culturally grounded.

As Núñez, Edwards, and Matos (1999) point out, "Meaning is in many ways socially constructed, but it is *not arbitrary*" (p. 53). The authors argue that meaning is subject to the constraints imposed by biological, embodied processes and the interaction of individuals within socially and culturally significant environments. Sfard (2009) also points out that the communicative effectiveness of gestures is strengthened by "our spontaneous ability, grounded in our cultural experience, to relate certain body movements to certain familiar things in the world" (p. 194). Indeed, the emerging theoretical perspective of embodied cognition posits that cognitive processes are deeply rooted in humans' bodily interactions with the surrounding world (Alibali & Nathan, 2012; Edwards, 2009; Núñez et al., 1999; Wilson, 2002). For instance, Alibali and Nathan suggested that the existence of gesture points to the framework of embodied cognition since a) gestures can reflect the grounding of cognition in the physical environment; b) representational gestures (those which are iconic and metaphoric) manifest mental simulations of action and perception; and c) metaphoric gestures reflect body-based conceptual metaphors. It is this additional connection to the body that influenced us to adopt embodied cognition as our theoretical perspective.

Methods

This research is part of a larger study where six experts and eight novices participated in a 90-minute audio and videotaped interview. The six experts included five mathematicians (Agustin, Ernie, Sarah, Luke, Diana) who frequently teach complex analysis or complex variables and one physicist (Bruno) who teaches courses that require notions from complex variables and who uses complex variables in his research. Novices were interviewed in four groups: two graduate pairs (Jim and BJ, Deb and Tim) and two undergraduate pairs (Gus and Rog, Amy and Kim). Participants were allowed as much time as needed to respond to various tasks, thus, some participants responded to more questions. The results of this study are based on three tasks, which all the participants completed. For each task the participants had access to a board with a figure of the Argand plane and two points z, in the second quadrant, and w, in the third quadrant pre-drawn. The participants were asked to use the diagram to complete the following tasks: 1) Where is $z+w$ located? 2) Where is zw located? and 3) Where is $1/z$ located? Participants were informed that we were interested in their geometrical interpretation of complex numbers including gestures, diagrams, illustrations, facial expressions, etc.

To best align with the goals of our research, we chose to define gesture as a movement of the upper limbs that fulfills a communicational function and is co-produced with mathematical speech. This definition is similar to definitions posed by others (Goldin-Meadow et al., 1999; Goldin-Meadow et al., 2001; McNeill, 1992). After transcribing all the interviews with descriptions of gestures, we began the coding process. Using our definition of gesture, all researchers concurrently coded one expert's interview, using McNeill's (2005) classification of gestures: *deictic*, *iconic*, *metaphoric*, or *beat*. McNeill described a deictic gesture as pointing with one's hand or fingers, an iconic gesture as a semantic image of a concrete entity, a metaphoric gesture as presenting an abstract image, and a beat gesture as a hand beating with time. An upper body motion was coded as a gesture if the movement occurred within two seconds of a participant's speech related to the current task.

After the initial classification gestures labeled iconic, deictic, or metaphoric were categorized as *dynamic* or *static* as defined by Marghetis and Núñez (2010). A dynamic gesture is a smooth and continuous motion, whereas a static gesture consists of beats and broken motion or a smooth movement bookended by beats or broken motion (p. 5). Gestures that had multiple phases, or multiple gestures within one phrase, and contained both static and dynamic segments were classified as *both* dynamic and static, whereas gestures with indistinct dynamic or static portions were classified as *ambiguous*. For both groups, gestures that were classified as either ambiguous or both made up approximately 10% of the total number of coded gestures (Experts: 10%; Novices: 10.5%) and were not included in the statistical analysis. Since beat gestures align with the prosody of speech and do not express semantic content, beat gestures were not classified as dynamic or static (Alibali & Nathan, 2012).

After the initial group coding session, five researchers individually transcribed and coded the remaining interviews. At least two researchers coded the same transcription in order to measure inter-rater reliability. Reliability between coders was established for both the number of gestures and number of dynamic and static gestures. The average inter-rater reliability (IRR) for number of gestures was 90.41%, while the average IRR for the number of static and dynamic gestures was 88.61%.

Results

An Analysis of Variance (ANOVA) model comprised of the factors Group (novice and expert), Task (addition, multiplication, and division), and Gesture (iconic, deictic, metaphoric, and beat) revealed that the factors Task and Gesture were statistically significant ($p < .001$ for each). The number of gestures exhibited for the multiplication task was significantly more than the number of gestures displayed for the addition task. Similarly the number of deictic gestures exhibited was higher compared to the other gestures. Group was, to our surprise, not a statistically significant factor ($p = 0.772$); this led us to perform a Hierarchical Cluster Analysis (HCA) which clusters the participants based on their gestures. In conducting the HCA, we explored the categories shown in Table 1 and each HCA was performed at the five-, four-, and three-cluster level. This analysis allowed us to explore how the participants were clustered for the different tasks and gestures. The HCA served as a guide in determining which participants would be compared and contrasted as part of the qualitative analysis.

Table 1

Cluster Categories

All Tasks by Gesture Type (All Tasks Type)	Division by Gesture Type (Div. Type)
All Tasks by Dynamic/Static (All Tasks D/S)	Division by Dynamic/Static (Div. D/S)
Addition by Gesture Type (Add. Type)	Iconic
Addition by Dynamic/Static (Add. D/S)	Deictic
Multiplication by Gesture Type (Mult. Type)	Metaphoric
Multiplication by Dynamic/Static (Mult. D/S)	Beat

A chief finding from the HCA was that one pair of novices, Amy and Kim, comprised their own cluster in six of the categories. Amy and Kim shared a cluster at all three levels with the male undergraduate students (Gus and Rog) for Division by Gesture type; this was one of few instances in which the novices gestured similarly. The graduate students (Jim and BJ, Tom and Deb) were clustered together in the Div. Type, All Tasks Type, and Deictic categories. A surprising result was that Gus and Rog shared a cluster with experts in all but three categories at the five-cluster level. Moreover, even within the expert category, participants did not always behave similarly. For instance, Bruno formed his own cluster in several categories, a finding we further substantiated qualitatively. The above clustering underscores the finding that Group was not a significant factor in the ANOVA model, despite the fact that novices were occasionally paired together.

The HCA corroborated the statistical significance of Task in our ANOVA. For instance, Bruno formed his own cluster in the Add. Type, Add. D/S, Mult. Type, and Mult. D/S categories. Additionally, Deb and Tom formed their own cluster at all cluster levels in the Add. Type and Add. D/S categories. Amy and Kim similarly formed their own cluster at all levels with respect to Mult. Type and Mult. D/S. Regarding the Gesture factor, the HCA did not suggest any particular trends amongst the participants. However, the percentages of different gesture types used by participants are useful in explaining the significance of Gesture as a factor. For instance, nearly three fourths of all gestures made by either of the graduate student pairs were deictic. In fact, the majority of gestures were deictic for all participants with the exception of Amy and Kim, who exhibited primarily iconic gestures (58.8%). Sarah also exhibited a large number of iconic gestures (40.9%). On the other hand, several participants displayed relatively few gestures in some categories. For example, less than ten percent of all gestures made by three novice pairs and one expert were metaphoric.

With a sample size of 10 interviews, differences between novices and experts may have been difficult to detect. Recognizing the substantial time commitment that coding interviews from a larger sample size would demand, we instead proceeded by conducting a qualitative extension of our quantitative findings, as suggested by Clement (2000). In order to tease out some of the intricacies not captured by our ANOVA but suggested by the HCA, we examined several episodes from the interviews through a qualitative lens. The qualitative component substantiated the lack of gestural homogeneity amongst our participants relative to their classification as a novice or expert that our quantitative results initially suggested. For example, during a particular portion of the multiplication task, novices Gus and Rog exhibited verbiage and iconic, dynamic gestures that closely resembled those of several experts (see Figures 1-3). Meanwhile, Amy and Kim demonstrated a rather unique conception of the complex arguments of z and w, treating one argument as a "negative angle" and the other as "positive" (see Figure 4).

Moreover, experts themselves didn't always gesture similarly, as evidenced by Bruno the physicist, who exhibited the largest proportion (32.4%) of metaphoric gestures out of all participants. Crucially, many of these metaphoric gestures coincided with unprompted verbiage discussing physics applications. Another significant finding was, unlike the novices, the experts consistently knew which form would be best for a particular task and effortlessly switched to another form. Furthermore, the novices with the longest interview times all struggled selecting an appropriate geometric representation, as well as switching between the algebraic and geometric representations of the tasks.

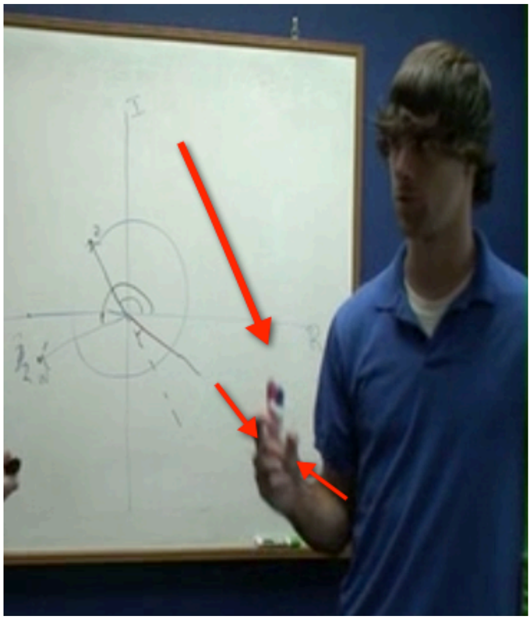


Figure 2. Rog: "Then we're going to be multiplying by factors less than one, so it [the modulus of zw] **is going to be shrinking**" (pinches his pointer finger and thumb, then continuously moves his hand downward).

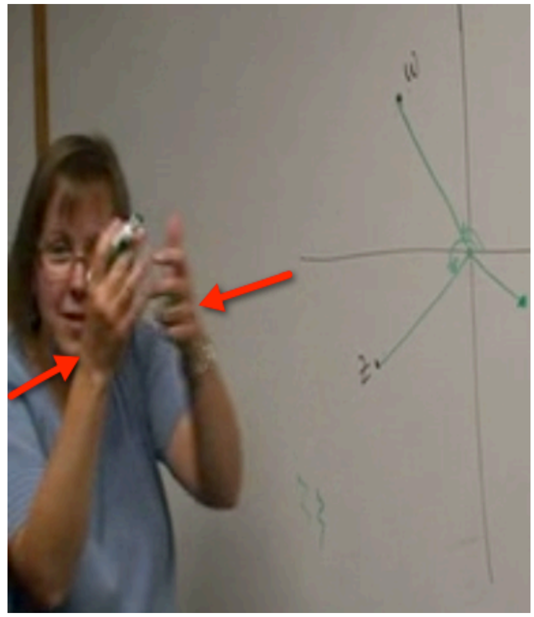


Figure 3. Sarah: "I'm still adding the angles, but now when I multiply the moduluses I'll get something smaller, right, because when you multiply two numbers less than one, **it** [the product] **gets even smaller**" (pinches hands together, palms facing each other).

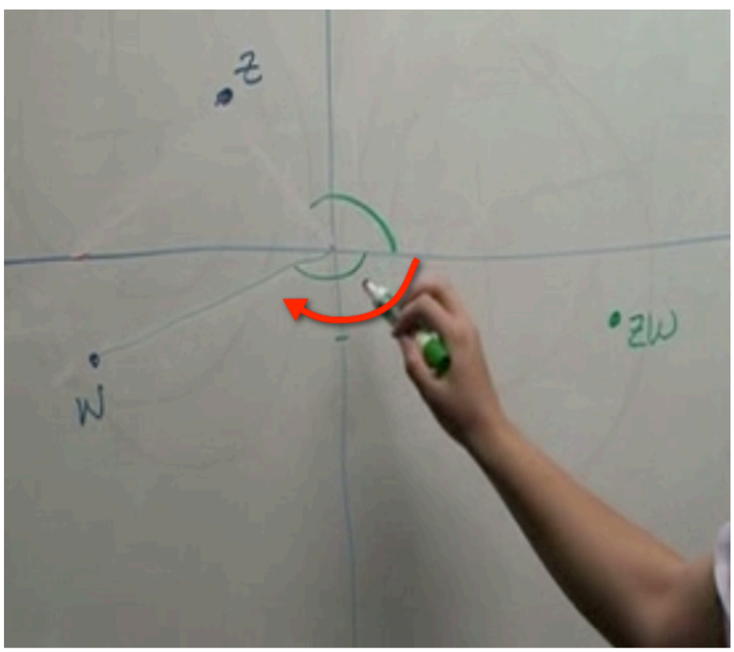


Figure 4. Kim depicted the argument of the complex number was a "negative angle," emanating clockwise from the positive real axis, whereas she chose the conventional orientation (counterclockwise from the positive real axis) for the argument of z.

Discussion and Implications

Our research underscores the conclusion of existing studies that possessing representational fluency is important for developing a comprehensive understanding of mathematics generally and complex arithmetic specifically (Danenhower, 2000; Panaoura et al., 2006; Zbiek, Heid, Blume & Dick, 2007). While most studies address geometric vs. algebraic representations, this study includes gestures as a representation and our results have implications for both teaching and further research.

The quantitative results indicate that both groups expressed more gestures for the multiplication task compared to the addition task. As such, it may be productive for instructors to acknowledge and highlight their own gestures when explaining the geometric interpretation of the product of two complex numbers. We also encourage instructors to have students explain their gestures. This may assist students to capture and to develop their own dynamic perspective as well as to help them become aware of the representation associated with their own gesturing. Another educational practice that has been linked to improving representational fluency is the incorporation of dynamic geometry programs such as *Geogebra*, *Geometer's Sketchpad,* or *Cabri Geometry* into the classroom. Such software has the potential to improve students' ability to conjecture, experiment, and devise novel solutions to problems, and have even been described as "cognitive reorganizers" (Arcavi & Hadas, 2000; Barrera-Mora & Reyes-Rodriguez, 2013; Olive, 2000; Robutti, 2006; Zbiek et al., 2007). While we naturally assumed the graduate students in our study would solve the tasks with greater ease and efficiency than the undergraduates, the results indicated the exact opposite outcome. Given that one of the main differences between the undergraduate and graduate student participants was the undergraduates' exposure to *Geogebra*, technology may account for the quantitative and qualitative differences. As such, conducting further research regarding the impact of gesture awareness and technology on students' conceptualization of complex numbers specifically could prove to be a fruitful endeavor for further research.

Given Bruno's gestures were quite different from other experts, another line of inquiry could explore the distinction of "mathematical expert." As a physicist, Bruno may employ his mathematical expertise in a unique way compared to mathematicians. This outcome suggests the possibility that gestural behavior may be significantly different depending on how a given expert is accustomed to using complex numbers in his or her area of expertise. Thus, a possible topic for future research could be exploring the existence of these differences amongst experts who use complex numbers in a variety of fields.

References

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science , 10* (4), 327-333.

Alibali, M., & Nathan, M. (2012). Embodiment in mathematics teaching and learning; evidence from learners' and teachers' gestures. *The Journal of the Learning Sciences , 21*, 247-286.

Arcavi, A., & Hadas, N. (2000). Computer mediated learning: an example of an approach. *International Journal of Computers for Mathematical Learning , 5*, 25-45.
Barrera-Mora, F., & Reyes-Rodriquez, A. (2013). Cognitive processes developed by students when solving mathematical problems within technological environments. *The Mathematics Enthusiast , 10* (1-2), 109-136.
Clement, J. (2000). Analysis of clinical interviews: foundations and model viability. In A. Kelly & R. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education*. Mahweh, NJ: Lawrence Erlbaum Associates, 547-589.
Danenhower, P. (2006). Introductory complex analysis at two British Columbia universities; the first week - complex numbers. *CBMS Issues in Mathematics Education* , 139-169.
Edwards, L. (2009). Gestures and conceptual integration in mathematical talk. *Educational Studies in Mathematics , 70*, 127-141.
Goldin-Meadow, S. (2003). *Hearing Gestures: How our hands help us think.* Chicago: University of Chicago Press.
Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology , 91* (4), 720-730.
Goldin-Meadow, S., Nusbaum, H., Kelly, S., & Wagner, S. (2001). Explaining math: gesturing lightens the load. *Psychological Science , 12* (6), 516-522.
Marghetis, T., & Nunez, R. (2010). Dynamic construals, static formalisms: evidence from co-speech gesture during mathematical proving. *Center for Research in Language Technical Report , 22* (1).
McNeill, D. (2005). *Gesture and Thought.* Chicago: The University of Chicago Press.
McNeill, D. (1992). *Hand and Mind: What gestures reveal about thought.* Chicago: The University of Chicago Press.
Miller, P. (2009). *Theories of developmental psychology.* New York: Worth.
Nemirovsky, R., Rasmussen, C., & Sweeney, G. W. (2012). When the classroom floor becomes the complex plane: addition and multiplication as ways of bodily navigation. *Journal of the Learning Sciences , 21* (2), 287-323.
Nunez, R., Edwards, L., & Matos, J. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational Studies in Mathematics , 39* (1-3), 45-65.
Olive, J. (2000). Implications of using dynamic geometry technology for teaching and learning. Plenary Paper presented at the. *Conference on Teaching and Learning Problems in Geometry* , Portugal, May 6-9.
Panaoura, A., Elia, I., Gagatis, A., & Giatilis, G.-P. (2006). Geometric and algebraic approaches in the concept of complex numbers. *International Journal of Mathematics Education in Science and Technology , 37* (6), 681-706.
Perry, M., Church, R., & Goldin-Meadow, S. (1988). Transitional knowledge in the aquisition of concepts. *Cognitive Development , 3*, 359-400.
Robutti, O. (2006). Motion, technology, gesture in interpreting graphs. *International Journal of Computer Algebra in Mathematics Education , 13*, 117-126.
Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification - the case of function. In E. Dubinsky & G. Harel (Eds.), MAA notes:

Vol 25. *The concept of function: aspects of epistemology and pedagogy*, Washington, DC: Mathematical Association of America, 59-84.

Sfard, A. (2009). What's all the fuss about gestures? A commentary. *Educational Studies in Mathematics , 70*, 191-200.

Sinclair, N., & Tabaghi, S. G. (2010). Drawing space: mathematicians' kinetic conceptions of eigenvectors. *Educational Studies in Mathematics , 74*, 223-240.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review , 9* (4), 625-636.

Zbiek, R. M., Heid, M. K., Blume, G. W., & Dick, T. P. (2007). Research on technology in mathematics education: a perspective of constructs. In F. K. Lester, *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1169-1207). Charlotte, NC: Information Age.

Understanding Students' Conceptualizations of Logical Tools

Casey Hawthorne

Abstract

While a significant amount of research has been devoted to exploring why university students struggle applying logic, limited work can be found on how students actually make sense of formal logic itself and the logical mechanisms used to communicate logical equivalence. This project borrows the theoretical framework of unitizing and reification, which have been effectively used to explain the types of integrated understanding required to make sense of symbols involved in numerical computation and algebraic manipulation, to investigate students' conceptualization of truth tables and implication statements. By using a continuum as a framework to analyze the degree to which students' thinking of each is compartmentalized versus unified, results indicate that students tend to favor one logical mechanism over another, without establishing a holistic view of both or an integrated view of the two together.

Introduction and Literature Review

As students continue to struggle bridging the gap from the more computationally focused mathematics courses of the K-12 US curriculum and low-level university courses to the more proof centered work in later mathematics studies, many universities have begun to incorporate the explicit instruction of proof into their curriculum, creating "transition to proof" courses (Moore, 1994). While the exact curriculum and coursework differs from institution to institution and class to class, a common component is the introduction of formal logic. For most mathematics educators its inclusion seems quite logical for multiple reasons. First of all, in order for students to write a proof, students must be aware of what is necessary to establish a statement is true or false (Epp, 2003). The rules that govern mathematical arguments and mathematical statements are different from those used in everyday speech. For example, the statements "There is a mother for all children" and "All children have a mother" are commonly used interchangeably inferring the later mathematical meaning (Dubinsky & Yiparaki, 2000). Teaching logic makes clear to students the approach and methods employed by the mathematical community. Also, logic creates a structure for students to shape and systematize their intuition. As students explore and develop examples, logic can help to organize these observations and illuminate their relevance as they apply for more general cases. It provides a supportive framework to actually teach students to reason mathematically. Logic acts as a scaffolding tool, making the abstract nature of argumentation more concrete (Epp, 2003).

As compelling as the instruction of logic appears, there is evidence that such a focus results in limited to no improvement in student achievement. In a well cited article Cheng, Holyoak, Nisbett, and Oliver (1986) found no difference in performance between university students who had taken an introductory logic course and a control group of students who had not. One explanation is that students have difficulty connecting abstract logic to real life applications. Selden & Selden (1995) found that students often struggle decoding statements written with a more familiar, colloquial structure and translating them into formal mathematical language. Unable to link the two domains, students are unable to tap into the power of formal logic.

To further examine the apparent divide between formal logic and its applications, this study was designed to explore the ways and extent in which students make use of formal logic to manipulate and work with conditional statements in different contexts. Individual student interview questions were purposely written in a recognizable if/then form in an effort to minimize translation issues and focus on how students coordinated the meaning between contextualized statements and corresponding symbolic representation(s). While the coordination between symbolic representations and contexts was the initial focus of the study, what emerged in the analysis was a different phenomenon. As students were asked to operate on symbolically written logical statements and comment on their associated understanding, their explanations revealed insight into not simply how they made use of the logic in relation to contexts, but to a much larger degree, how they thought about and made sense of the logical tools themselves, in particular the symbolic expression $p \rightarrow q$ and its related truth table. As participants employed these logical tools, it became clear that some students viewed the hypothesis p and conclusion q as quite separate entities, while others seemed to considered the conditional statement in a more holistic way. Similarly, in comparing the logical equivalence of various statement forms ($p \rightarrow q$, $\sim p \vee q$, $\sim q \rightarrow \sim p$), some students treated the different cases of possible truth values for p and q in a piecewise or compartmentalized manner, while others conceptualized the various combinations as more connected and unified. Therefore, it seems suitable to use these views to represent different ends of a continuum as a means to frame how students conceptualize logical tools.

Theoretical Background

To date, math educators have put forward various characterizations to capture the integrated and hierarchal nature of mathematics and more importantly the type of knowledge necessary to fully conceptualize it. One earlier such effort was Steffe (1983) who proposed the notion of unitizing in the context of numbers. He asserted that for students to possess a rich understanding of numbers, they cannot be presented as some pre-existing entities that students find and retrieve, but rather must be envisioned as composite unities that are actively constructed as a collection of individuals taken together. Lamon (1996) added to this notion, describing unitizing more generally as the cognitive assignment of multiple mathematical entities into a combined whole. This newly constructed abstract object can then be used to reason with as a single unit. Lamon emphasized that the key to unitizing is the ability to connect multiple ideas and envision them as a single, collective chunk while at the same time retaining an appreciation of their individual parts relative to each other and to the newly conceptualized whole.

Another theoretical framework that captures this phenomenon is the idea of reification. Mathematical constructs can be understood not only at different levels of their embedded hierarchal structure, but also in terms of the different roles they play. Depending on one's point of view, a given mathematical concept, while expressed by a single symbolic representation, can simultaneously take on two distinct, yet connected interpretations. As such, mathematical notation naturally possesses a double meaning. First, mathematical symbols can take on an *operational conception,* where notation is viewed as a set of instructions for a particular process (Sfard & Linchevski, 1994). At the same time, these expressions can embody a *structural conception* and represent the result of these processes. With this interpretation, the process has been reified and each of the various computations is considered as a whole unified and completed object. As such, this newly created mathematical entity can then be treated as a single unit and acted on as an object. In this way, the mathematical object becomes the basis for a new process, resulting each time in a more complicated and tightly packed symbol.

As Sfard (1995) points out, this dual role enables mathematical notation to be an extremely powerful tool. It allows the user to understand and conceptualize a very complex and involved process, while at the same time treat it as a simple object, disconnected from semantic meaning. It can then be manipulated and simplified syntactically, without the large burden that the operational mode of thinking places on the working memory. This capacity though can also act as a double edged sword, what Sfard and Linchevski (1994) refer to as a pseudostructural conceptualization. Often students are introduced to powerful symbolic notation along with various procedures to apply to them, but fail to develop an underlying grasp of the processes the notation embodies. Students with such an approach might be able to consider the symbols as objects, what seems like reification, but the notation is void of any semantic meaning. As such, they are unable to connect any conceptual meaning to the notation and simply carry out algorithms mechanically, without any understanding of the significance of their actions. With the operational underpinnings of the abstract objects far removed, they are unable to deviate from the systematic techniques and flexibly apply any associated procedures. As Sfard (1995) highlights, notational expressions become viewed as "meaningless symbols governed by arbitrary established transformations" (p. 30). In the end, the manipulation itself becomes the focus of the activity, and the symbolic results are seen as producing the answer themselves.

While these two constructs, unitizing and reification, have been used to explain the types of integrated understanding required for numerical computation and algebraic manipulation, a similar conceptualization seems to be necessary for the symbolic understanding and application of logic as well. As students compartmentalize different notational pieces of logical symbols, or treat the symbols without reference to the semantics they represent, it seems evident that they are failing to appreciate the different layers and meanings which the symbols embody. Therefore, this study makes use of these theoretical constructs as a basis to understand students' ability to conceptualize various notational structures used in formal logic, namely implication statements and associated logic tables, as they make sense of logical equivalence.

Methods

The participants for this study were all drawn from a discrete mathematics course that simultaneously functions as a transition to proof course, introducing and engaging students in the fundamental elements of mathematical proof and communication, and as a computer science course, ensuring coverage of the mathematical content ordinarily associated with discrete mathematics such as set theory, logic, induction, and combinatorics. The six participants, interviewed towards the end of the course, were all in the top quarter of the class, having demonstrated mastery of topics on assessments throughout the semester. Each student participated in a 60-minute semi-structured clinical interview (Ginsburg, 1997). As such, a detailed protocol was used to guide the interview, but the researcher followed up with multiple clarifying questions to develop a more detailed understanding of each participant's precise thinking. While the interview protocol consisted of two main sections, almost all of the relative data were taken from the first half which was designed to explore how participants interpreted and made sense of notation and logical statements in symbolic form. This section comprised of two main questions. First, the participants were asked to analyze the equivalence of various logical statements relative to $p \rightarrow q$ and explain how they understood such equivalence or not. Second, they were asked to make sense of the negation of a conditional statement, presented in the symbolic form $\neg(p \rightarrow q)$ and give an example to illustrate their understanding.

Each interview was videotaped and transcribed. The participants' answers were reviewed using a grounded theory approach (Strauss & Corbin, 1994). The initial coding pass relied on open coding in which evidence was collected to make sense of how students conceptualized logical statements and interpreted logical equivalence. After a detailed review of the videos and their accompanying transcripts, data suggested that it was the degree to which participants viewed both logical statements in general and their associated logic tables in a unified or compartmentalized fashion that distinguished the participants' approaches to making sense of logical equivalence. A continuum was proposed along these two dimensions (statements and logic tables) as to the extent they were considered holistically or not and each participants' answers were further analyzed and compared in terms of these two scales. After the student interviews were initially examined and reviewed, the professor of the course was then asked the same questions in order to provide a comparison between the students' conceptualizations to that of an expert.

Results

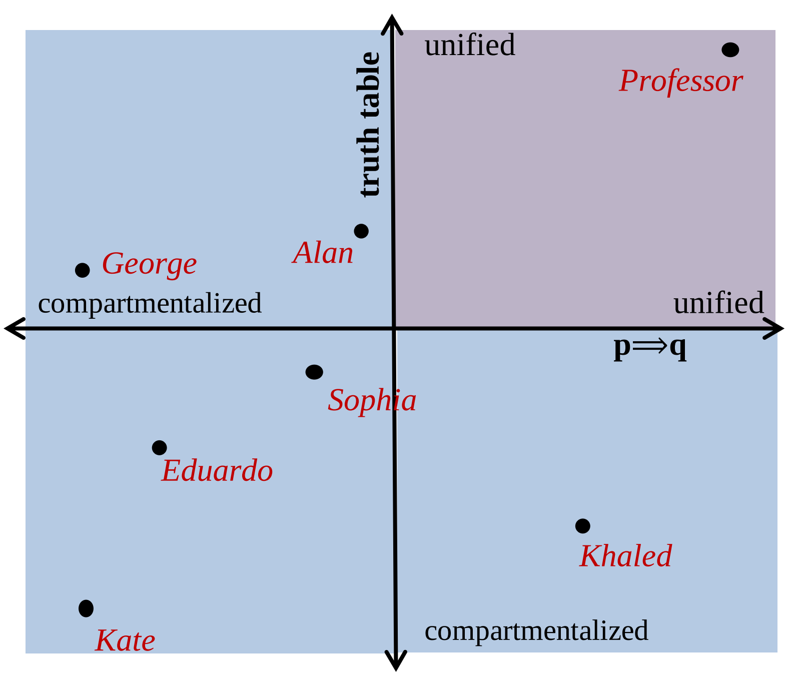
Using a continuum as a framework to analyze the degree to which students' conceptualizations of both truth tables and implication statements are compartmentalized versus unified provided interesting insight into how students make sense of these logical mechanisms. As highlighted in the following chart, the analysis resulted in specific descriptions in the various ends of each spectrum. Students who tended to compartmentalize truth tables, often did not fully appreciate that the given compound statements could have four possible truth combinations. They translated the symbols directly, interpreting $\neg p \wedge q$ to mean that p is false and q is true. In other cases they demonstrated a belief that the symbols took on only the meaning that would make the statement true, for example, $p \rightarrow q$ was construed to mean p is true and q is true. In addition, once constructed, these students did not compare statements by the four cases collectively. Often, they believed that only one line of the truth table corresponding between two statements was sufficient to establish equivalence. This was in comparison to students who had a unified conception of the truth tables. Most of these students were able to quickly run through the various possibilities in their head. They strategically focused on cases that they knew might not be equivalent. When they produced a truth table, it was as a way to record their previously established verbal ideas and better organize their thinking.

<table>
<tr><th colspan="2">Truth Table</th></tr>
<tr><th>Compartmentalized</th><th>Unified</th></tr>
<tr><td>• Interpreting symbols as only taking on the meaning that makes the statement true. (For example, p→q means p is true and q is true)
• Considering symbols literally, or as a direct translation. (For example: ¬p∧q means p is false and q is true.)</td><td>• Seeing the truth table as an organizational tool to represent the four possible cases
• Ability to consolidate cases when strategically useful
• Connecting semantic meaning to each case
• Facility in determining missing parts given the truth of any of the various components or of a compound statement as a whole.</td></tr>
</table>

For the students who compartmentalized the implication statement $p \rightarrow q$, their approach to compare equivalence or negate statements tended to involve a method similar to algebra as a means to manipulate symbols. Not knowing any rules to negate an implication, they would substitute the "or" statement $\neg p \vee q$ without any explanation why this was possible in order to proceed. Even when given or when they provided an explicit real life example, they continued to make use of symbols to determine valid changes and then translate the final results back into words. Interestingly, not one student who successfully arrived symbolically at the equivalence $(p \rightarrow q) \leftrightarrow (\neg p \vee q)$, was able to explain semantically why these were equivalent. In contrast, students who indicated a unified understanding of $p \rightarrow q$ used an example to explain why the negation would be p and not q. Unfortunately, both students who used this approach were unable to correctly express their understood equivalence symbolically. One (Sofia) resorted back to what she believed would be a rational symbolic way, negating each piece separately for $\neg p \rightarrow \neg q$. The other (Khaled) tried to simply translate his final idea into symbols, writing $p \rightarrow \neg q$, indicating still a compartmentalized appreciation of the symbols.

$p \rightarrow q$	
Compartmentalized	unified
• Looks at symbols in an isolated manner without making meaning of them as a whole piece. (For example: negates conditional statement by distributing negation to both p and q.) • Compares statements based on notational structure instead of meaning. (For example: negation of implication must be an implication.) • Tendency to invoke algebraic symbol manipulation	• Comprehending the premise and conclusion as one statement. • Connecting the symbolic representation to its semantic meaning.

Finally, after each student was assessed on each continuum separately, their results were graphed on an axis containing both continuums simultaneous as shown below. This provided an interesting analysis as it highlights that no student demonstrated a unified conception of both logical structures. While almost every student possessed a unified conception of one, the two seemed almost inversely proportional. Also, students seemed to compensate for which ever conceptualization was compartmentalized by attempting to use the other mechanism, even when it was not productive. This was in contrast to the professor who demonstrated a clearly unified conception of each logical structure. Not only did he articulate a unified understanding for both, but in his description, he was unable to disentangle the idea of an implication from its associated truth table. He quickly alternated between discussing $p \rightarrow q$ holistically, then in parts, then in terms of its various possible truth values.



Discussion

This paper has both potential research contributions as well as instructional implications. As for research, this project explores a seemingly overlooked area of logic. While a significant amount of research has been done to understand students' struggles in applying logic, very little work has gone into investigating students' conceptualization of the notational tools used to understand logical equivalence. By bringing in constructs from other areas of mathematics, it provides tools for researchers to analyze and describe the type of understanding necessary to make sense of the symbols associated with logic. Further research could be devoted to looking into connections between where students fall on these continuums and their ability to flexibly apply logic in different applications. It seems reasonable that a narrow view of the underlying logical mechanisms could contribute significantly to a student's ability to apply them.

In terms of instructional implications, this work provides new insight into students' conceptualization of logical structures and associated difficulties. As such, these findings have the potential to make professors of discrete math or transitional courses more aware of the type student thinking connected with the logical structures presented in class. It provides them with specific areas to attend to with their students and break down what might previously be categorized as a general problem with logic into concrete difficulties.

References

Barnard, T. (1995). The impact of meaning on students' ability to negate statements. In *PME conference* (Vol. 2, pp. 2-3). The program committee of the 19th PME conference.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive psychology*, 18(3), 293-328.

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. Schoenfeld & J. Kaput (Eds.), *Research in Collegiate Mathematics Education IV - Issues in Mathematics Education* (Vol. IV): CB MS.

Epp, S. S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly*, *110*(10), 886-899.

Ginsburg, H. (1997). *Entering the child's mind*. Cambridge, UK; Cambridge University Press.

Lamon, S. J. (1996). The development of unitizing: Its role in children's partitioning strategies. *Journal for Research in Mathematics Education*, 170-193.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, 27(3), 249-266.

Selden, J. and A. Selden (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29, 123-151.

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification—the case of algebra. In *Learning mathematics* (pp. 87-124). Springer Netherlands.

Sfard, A. "The development of algebra: Confronting historical and psychological perspectives." *The Journal of Mathematical Behavior* 14.1 (1995): 15-39.

Steffe, L. P. (1983). Children's algorithms as schemes. *Educational Studies in Mathematics*, 14(2), 109-125.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273 – 285). Thousand Oaks: Sage Publications.

Key Words: Logic Table, Discrete Mathematics, Logical Equivalence

EVALUATING PROFESSIONAL DEVELOPMENT WORKSHOPS QUICKLY AND EFFECTIVELY

Charles Hayward and Sandra Laursen
Ethnography & Evaluation Research, University of Colorado Boulder

Abstract

Many funding agencies require evaluation of the impact of professional development projects they support. However, improved student outcomes, the ultimate goal, may take longer to be realized than the project time frame allows. Instructors need time to implement and refine new skills before positive student outcomes are realized, a delay that may be exacerbated in classes that are not taught frequently. We report on one example of an efficient and cost-effective self-report measure designed to detect the initial changes in teaching practices that lead to improved student outcomes over time. We discuss the ability for timely and accurate measures through this instrument. Results support the interpretation that instructors' reported teaching practices show changes consistent with methods taught at professional development workshops on Inquiry-Based Learning in mathematics. Additionally, correlations with self-reported level of implementation suggest that instructors are reporting honestly, and not just socially desirable changes consistent with their concept of "real Inquiry-Based Learning."

Key words: Evaluation, Measurement, Professional Development, Inquiry-Based Learning

Background

After decades of innovation and research it is clear that certain reforms of classroom practice improve undergraduate education in science, technology, engineering and mathematics (STEM). Research in cognitive science and education offers persuasive evidence that students can and do learn better through active, student-centered forms of instruction (Hake, 1998; Springer, Stanne, & Donovan, 1999; Bransford, Brown, & Cocking, 2000; Prince, 2004; Ruiz-Primo, Briggs, Iverson, Talbot, & Shepard, 2011; Deslauriers, Schelew, & Wieman, 2011). Yet relatively few students experience these proven, "high-impact" educational practices during college (Kuh, 2008). The President's Council of Advisors on Science and Technology (PCAST, 2012) advocated active learning strategies in order to meet its goal of an additional 1 million STEM graduates over the next decade.

Uptake of these methods by large numbers of faculty at diverse institutions is now the bottleneck in improving STEM higher education (Fairweather, 2008; Henderson & Dancy, 2007; 2008; 2011). To broaden uptake of student-centered teaching and learning approaches, professional development of college instructors (CIPD) is crucial. But efforts to broaden the reach of CIPD must be based on good evaluation evidence about whether it is having the desired effect on teaching. While improved student learning is the ultimate goal of CIPD, measuring student outcomes directly is not always feasible due to the cost and complexity. Additionally, the impact on students lies far downstream from the intervention itself (Guskey, 2002) as instructors must apply and refine the methods before positive student results can be detected. Instructors may not teach a course every year, resulting in a large lag between the CIPD and any detection of positive student outcomes.

Given this time lag, positive student outcome results may help to demonstrate the merit of a particular CIPD intervention, but they do not provide formative feedback to help instructor developers to diagnose or improve a particular intervention. Measuring the impact of CIPD through its effect on student outcomes can contribute to the research base, but does not provide a nuanced and flexible evaluation tool that is responsive on the time scale on which

CI developers—those who plan and lead the professional development—must adapt, refine, and report results of their intervention. Rather, we need reliable and valid evaluation tools to measure whether and to what extent CIs have made changes to their instruction as a result of CIPD interventions. Measuring changes in instructional practices is challenging. Multiple observation protocols have been developed to measure changes in teaching practices, however they are often time-consuming and disruptive to classrooms (Hora & Ferrare, 2012). Surveys are easier to conduct, but some have argued that they are inaccurate due to respondent biases (Desimone, 2009).

This study reports on evaluation of professional development workshops for college instructors on inquiry-based learning (IBL) in mathematics. IBL is a form of active learning that helps students develop critical thinking through ill-defined problems and by constructing and evaluating mathematical arguments. IBL is based on the teaching practices of R.L. Moore (1882-1974), a mathematician at the University of Texas, Austin. His teaching method involved students using definitions, logic, and precise language to prove mathematical theorems (Mahavier, 1999). Students worked independently and were not allowed to consult other students or textbooks. They then presented proofs in class and were critiqued. Today, this method is typically modified to allow more student collaboration and is referred to as the Modified Moore Method. IBL has emerged as a broader umbrella term encompassing Moore's method as well as others that share the spirit of student inquiry through deep engagement with mathematics and collaboration with peers (Yoshinobu & Jones, 2013). [For an example, see (Schumacher, 2010).] In all forms of IBL, students learn through analyzing ill-defined problems and constructing and evaluating arguments (Prince & Felder, 2007; Savin-Baden & Major, 2004). This supports deep learning of mathematical concepts (Moon, 2004; McCann, Johannessen, Kahn, & Smagorinsky, 2004). To teach in this manner, many instructors must transition from traditional lecture methods to more student-centered teaching approaches. [For an example, see (Retsek, 2013).] The professional development workshops we have studied aim to help instructors make this transition.

Conceptual Framework

Guskey (2000) classifies evaluation of professional development (PD) into five levels of increasing complexity. At higher levels, evaluation requires increased time and resources. Each level builds upon those before it and varies as to the questions evaluators address, how the data is gathered, what is measured, and how the information is used. The first level comprises participants' immediate reactions to the PD, while Level 2 goes further to address what participants have learned from the PD. In Level 3, evaluation measures organizational support and change. Participants' use of new knowledge and skills is measured in Level 4, and Level 5, the most complex, addresses student learning outcomes. Our larger project evaluates professional development workshops for college instructors on Inquiry-Based Learning in mathematics at Levels 1 through 5. Since student learning (Level 5) is very difficult to assess within the timeframe of grant-funded projects, this report focuses on the next highest level (Level 4), evaluation of participants' use of new knowledge and skills.

While many evaluation efforts at this level use classroom observation protocols to assess participants' implementation of methods presented during professional development workshops, these are time- and resource-intensive and may interfere with normal classroom dynamics (Guskey, 2000). On the other hand, surveys are cost-effective but rely on self-report, which may be prone to bias. Participants are not always good at judging their own learning since they do not yet have accurate benchmarks (Kruger & Dunning, 1999). Self-report may also be affected by social desirability if participants feel pressure to answer a certain way (Desimone, 2009). However, when instructors report concrete behaviors without evaluative components, self-reports correspond well with observations (Desimone, 2009).

Therefore, our instrument minimizes these self-report pitfalls by having teachers report on their use of concrete teaching practices, rather than subjectively judging their own knowledge or evaluating the quality with which they implement new techniques.

While evaluators should be concerned with the quality of implementation, this will likely improve over time and may require repeated measurements. Many professional development programs are funded through short-term grants. Time is spent developing and conducting the professional development, and therefore, not much time is left for participants to use and develop the new knowledge and skills in their own classrooms before evaluation must be conducted. As a result, professional development workshops may be deemed ineffective when in reality, skills may continue to develop and benefits may not be fully realized until well after the project has ended.

In this paper, we report on a self-report measure for IBL workshops that is designed to quickly and accurately detect the initial changes in teaching practices following professional development workshops. Our main research questions are:

1) Are cost-effective and efficient self-report measures of changes in teaching practices an accurate way to evaluate outcomes of professional development?
2) In what capacity can evaluation efforts assess the implementation of new knowledge and skills from professional development workshops within the short-term cycle of grant funding?

Research Methods

The study sites were four workshops for mathematics faculty, led by universities with IBL Mathematics Centers where an extensive menu of IBL courses had been developed and taught over several years. Thus, faculty with expertise on IBL were available to lead each workshop. Through funding from the National Science Foundation, the universities developed and implemented annual IBL workshops from 2010 to 2013 for four cohorts of math faculty new to IBL. Workshops spanned four or five days and included a mix of invited talks, open discussions, video observations, expert panels, hands-on exercises, and work time. Each workshop had a slightly different style; the 2010 and 2012 workshops were highly interactive, while the 2011 and 2013 workshops were more conference-like.

As evaluators for the workshop project, our team conducted pre- and post- workshop surveys at each workshop. We also conducted one-year follow-up surveys for the first three cohorts (2010 through 2012). All three surveys included both quantitative items and open-ended questions. Evaluation instruments addressed Levels 1-5 in Guskey's framework. Level 1 was assessed on post-workshop surveys where participants rated and commented on the quality of the workshop and logistics and the aspects they found most and least helpful. Level 2 was measured with Likert-scale items to reflect participants' knowledge, skills, and beliefs about inquiry teaching, as well as their motivation to use inquiry methods. By assessing these items before participants attended the workshop, immediately afterwards and again one year later, we could identify significant changes in their knowledge and perceptions. Participants also wrote definitions of IBL on each survey to reveal their current perception and level of understanding. To assess Level 3, participants rated the levels of support for IBL from their departments, their chairs, their colleagues, and their institutions. Participants also completed open-ended responses about ways they have and have not been supported in implementing IBL. On follow-up surveys, thirteen Likert-scale items and two open-ended items addressed student gains (Level 5) from IBL.

A large portion of the follow-up survey was aimed at Level 4 evaluation. In one item, participants reported if they had not implemented IBL techniques, implemented some IBL techniques, or implemented one or more fully-IBL courses. Teachers also rated the frequency with which they used eleven different teaching practices. Both pre-workshop and follow-up surveys included these eleven items, so comparisons allowed us to detect changes for each individual instructor. Some items described teaching behaviors consistent with IBL methods while other items were not. Open-ended questions collected data on the challenges and supports experienced in implementing IBL techniques in the first year after the workshop.

Results

Participants

In total, 167 participants attended the workshops. They came from a variety of institutions. Most taught at four-year colleges (37%), Ph.D.-granting research universities (37%), or master's-granting comprehensive universities (23%), and a small number taught at two-year colleges (4%). About 13% of participants taught at minority-serving institutions. Many were tenure-track faculty that were not yet tenured (35%); some were tenured (34%), and some were not tenure-track (27%). A small number were high school teachers (<1%) and graduate students (3%). The largest group had between 2 and 5 years of teaching experience (27%), while some had less than 2 years experience (20%), and many had more: 19% had 6-10 years experience, 18% had 11-20 years experience, and 16% had more than 20 years experience. A small number had experienced IBL classes as a student (25%) and almost half had some prior experience using IBL methods as an instructor (46%).

Most participants were male (58%), but the percentage of women (42%) was higher than among math faculty as a whole (National Science Foundation, 2008). Most participants identified as of European descent (74%), and a small percentage were of Asian descent (10%). These proportions are about the same as in U.S. math faculty as a whole (National Science Foundation, 2008).

While pre-workshop, post-workshop, and follow-up surveys were all collected anonymously, they were matched using two pieces of non-identifying individual information. Details about the numbers of surveys collected from each workshop are presented in Table 1.

Table 1. Survey response rates as a percentage of attendees.

Cohort	Attendees	Pre-Surveys	Post-Surveys	Matched Pre/Post	Follow-Up	Matched Pre/FU
2010	42	37 (88%)	41 (98%)	33 (79%)	31 (74%)	23 (55%)
2011	55	47 (85%)	43 (78%)	29 (52%)	36 (65%)	21 (38%)
2012	42	40 (93%)	41 (95%)	38 (88%)	29 (69%)	25 (60%)
2013	28	22 (79%)	26 (93%)	21 (75%)	*TBD, Fall 2014*	
Total	167	146 (87%)	151 (90%)	121 (72%)	96 (69%)	69 (50%)

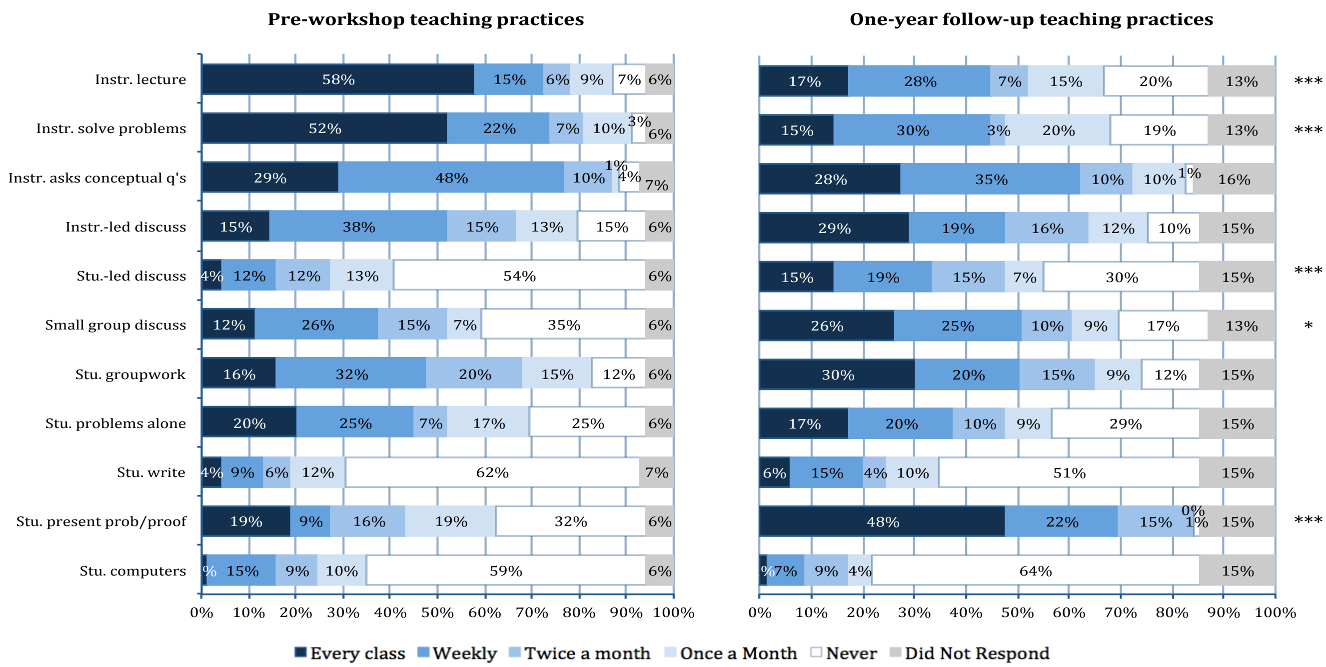
Teaching Practices: Changes in Practices

Participants reported teaching practices on both pre-workshop surveys and follow-up surveys. For each specific practice, respondents indicated on a 5-point scale whether they did it in 'every class' (5), 'weekly' (4), 'Twice a month' (3), 'Once a month' (2), or 'Never' (1). On follow-up surveys for the first three cohorts, 69% of participants responded, of which 69 (50% of attendees) supplied matching pre-workshop and follow-up surveys. We used Wilcoxon Signed Rank tests to measure changes in the ordinal data.

Teachers reported changes in practices consistent with IBL teaching. Overall, participants reported significant decreases in the frequencies of *Instructors lecturing,* and *Instructors solving problems at the board.* They also reported significant increases in *Student-led whole class discussions, Small group discussions,* and *Students presenting problems or proofs.*

There were no significant differences in practices that are not specific to IBL methods, including *Instructors asking conceptual questions, Instructor-led whole class discussions, Students solving problems individually, Students writing individually, or Computer assisted learning.* These non-IBL items were added to detect response biases; non-significant changes suggest that instructors are using the full scales and are reporting honestly. Indeed, in our other evaluation of professional developments, we have found that college instructors tend to be more self-critical on surveys and use the full range of scales, whereas K-12 teachers tend to use just the extreme answers (Hayward, Laursen, & Thiry, 2013). Additionally, *Student collaborative work in small groups* is characteristic of some types of IBL teaching. While participants did report increased use of this strategy, the difference was outside the range of statistical significance (p=0.086).Teaching practices on pre-workshop surveys and one-year follow-up surveys are compared below in Figure 1.

Figure 1. Self-reported teaching practices.

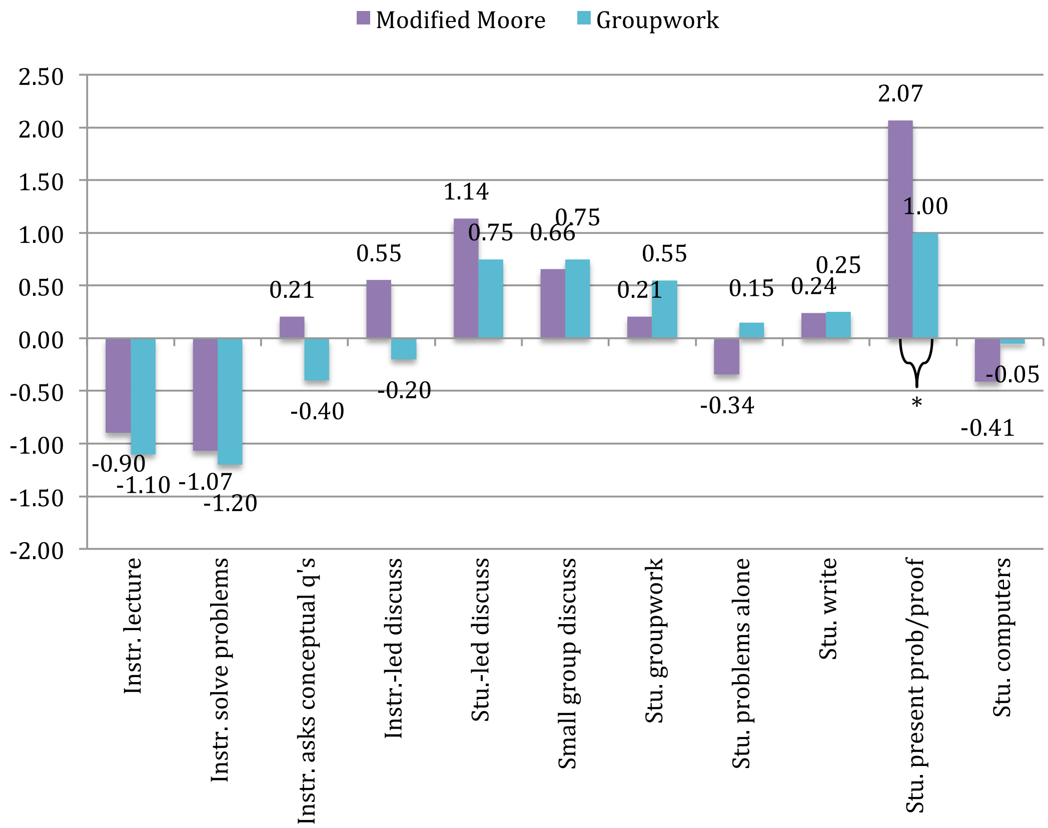


*p<0.05, **p<0.01, ***p<0.001

Teaching practices data were also analyzed for differences between two different types of IBL presentations. All participants in the 2011 workshop and about half of the participants in the 2012 workshop were presented with a groupwork-centered version of IBL (n=20 respondents), while all participants at the 2010 workshop and the other half of participants in the 2012 workshop were presented with a Modified Moore-method approach to IBL that uses individual student presentation of proofs or solutions followed by class discussions (n= 29 respondents).

When slicing up the data this much, sample sizes were too small to detect significant differences for most items, but the data suggest that participants presented with the Modified Moore method reported greater increases in the frequencies of having students present problems or proofs compared to participants presented with a groupwork-centered version of IBL. These changes are detailed in Figure 2.

Figure 2. Changes in teaching practices by cohort.



$^*p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Taken together, these findings suggest that our self-report instrument may be sensitive to differences in workshops. Practices that are common to both versions of IBL, including decreases in instructors lecturing and solving problems, showed significant changes regardless of the type of IBL presented. However, participants who were presented with a Modified Moore Method reported increases that were statistically greater than those participants presented with groupwork-centered IBL.

Teaching Practices: Accuracy of Self-Report

In order to assess accuracy of self-reported implementation of IBL techniques, two approaches to probing participants' use of IBL were compared. Participants reported in a single item on the follow-up survey whether they had not implemented IBL, implemented some elements of IBL, or fully implemented IBL in one or more courses. These were converted to numerical answers of 1, 2, or 3, respectively. This single item was compared to the eleven separate items on specific teaching practices.

The results showed significant negative correlations between implementation level and frequencies of *Instructors lecturing* ($r = -0.62$, $p < 0.001$), *Instructors solving problems at the board* ($r = -0.49$, $p < .001$), *Instructors asking conceptual questions* ($r = -0.46$, $p < 0.001$), and *Students solving problems individually* ($r = -0.53$, $p < 0.001$). The only significant positive correlation was with *Students presenting problems or proofs* ($r = 0.39$, $p < 0.01$).

These features are consistent with a definition of IBL aligned with that of its founder, R.L. Moore, in which the instructor takes a less active role and class time is largely composed of students presenting and critiquing proofs that they have worked on before class (Jones, 1977). So, while participants seem to have self-identified their one-year follow-up teaching practices by using the Moore Method as a benchmark for "real IBL," the changes in their teaching practices in comparison to pre-workshop surveys encompass a broader definition of IBL inclusive of both whole-class and small group discussions. Since instructors have reported teaching practices different than those they consider to be "real IBL," this suggests that participants are providing honest responses rather than the socially desirable responses.

Implications for Future Research

Initial results suggest that in addition to being cost-effective and efficient, this self-report measure of changes in teaching practice shows promising indicators of accuracy. Changes from pre-workshop surveys to one-year follow-up surveys were consistent with the teaching practices presented at the IBL workshops. Trends in data suggest that self-report measures may also be sensitive to the type of IBL presented. Correlations on one-year follow-up surveys between self-reported implementation levels of IBL and teaching practices indicate that participants largely consider the Moore Method to be the "real" IBL. So, participants are reporting changes in their teaching practices in line with those presented at the workshop, but they do not necessarily identify themselves as doing "real IBL." One of the main critiques of the accuracy of self-report measures is that participants often report only socially desirable answers (Desimone, 2009). However, these nuanced differences indicate that for this measure, instructors may be accurately reporting teaching practices that are not consistent with the socially desirable definition of "real IBL." These surveys will continue to be used to evaluate future workshops, and increased sample sizes should provide greater statistical power. Additionally, we have received NSF funding to formally validate the survey instrument through comparisons with classroom observations.

This measure was simple and efficient to administer. While we did not address the quality of implementation, we were able to measure the extent of change in practice following the professional development workshops in a timely and cost-effective manner. This self-report tool is especially useful given the short time frame and tight budgets of many grant-funded PD projects. While this particular survey is best suited for IBL workshops, it could easily be adapted to professional development workshops on other topics or in other disciplines by changing the target instructional practices.

References

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school.* Washington D.C.: National Academy Press.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Education Researcher , 38* (3), 181-199.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science , 332* (6031), 862-864.

Fairweather, J. (2008). *Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education: A status report for the National Academies Research Council Board of Science Education.* Retrieved April 23, 2011, from http://www7.nationalacademies.org/bose/Fairweather_CommissionedPaper.pdf

Guskey, T. R. (2000). *Evaluating professional development.* Thousand Oaks, CA: Crowin Press.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice , 8* (3/4), 381-391.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics , 66* (1), 64-74.

Hayward, C., Laursen, S., & Thiry, H. (2013). *Survey of teacher participants in professional development workshops conducted by the Biological Sciences Initiative, 2010 – 2012: Final report.* Report to the Biological Sciences Initiative. Boulder, CO: Ethnography & Evaluation Research.

Henderson, C., & Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics - Physics Education Reform , 3* (2), 0201102.

Henderson, C., & Dancy, M. H. (2011). *Increasing the impact and diffusion of STEM education innovations.* Retrieved July 1, 2013, from Comissioned paper for Forum on Characterizing the Impace and Diffusion of Engineering Education Innovations: http://www.nae.edu/File.aspx?id=36304

Henderson, C., & Dancy, M. H. (2008). Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations. *American Journal of Physics , 76*, 79-91.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher , 30* (March), 141-158.

Hora, M. T., & Ferrare, J. J. (2012). *A review of classroom observation techniques used in postsecondary settings.* White paper. Prepared for the Measurement of Teaching Practices in Undergraduate STEM workshop hosted by AAAS/NSF.

Jones, F. B. (1977). The Moore method. *The American Mathematical Monthly , 84* (4), 273-278.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology , 77* (6), 1121-1134.

Kuh, G. (2008). *High-impace educational practices: What they are, who has access to them, and why they matter.* Washington D.C.: AAC&U.

Libarkin, J., & Anserson, S. W. (2005). Assessment of learning in entry-level geoscience courses: Results from the Geoscience Concept Inventory. *Journal of Geoscience Education , 53*, 394-401.

Mahavier, W. S. (1999). What is the Moore Method? *Problems, Resources, and Issues in Mathematics Undergraduate Studies , 9* (4), 339-354.

McCann, T. M., Johannessen, L. R., Kahn, R., & Smagorinsky, P. (Eds.). (2004). *Reflective teaching, reflective learning: How to develop critically engaged readers, writers, and spearkers.* Portsmouth, NH: Heinemann.

Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice.* New York: RoutledgeFalmer.

National Science Foundation. (2008). *TABLE 3. Employed doctoral scientists and engineers in 4-year educational institutions, by broad field of doctorate, sex, faculty rank, and years since doctorate: 2008*. Retrieved July 29, 2013, from Characteristics of Doctoral Scientists and Engineers in the United States: 2008: http://www.nsf.gov/statistics/nsf13302/pdf/tab3.pdf

National Science Foundation. (2008). *TABLE 4. Employed doctoral scientists and engineers, by selected demographic characteristics and broad field of doctorate: 2008*. Retrieved July 29, 2013, from Characteristics of Doctoral Scientists and Engineers in the United States: 2008: http://www.nsf.gov/statistics/nsf13302/pdf/tab4.pdf

PCAST. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Executive Office of the President.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education , 93* (3), 223-231.

Prince, M., & Felder, R. (2007). The many facets of inductive teaching and learning. *Journal of College Science Teaching , 36* (5), 14-20.

Retsek, D. Q. (2013). Chop wood, carry water, use definitions: Survival lessons of an IBL rookie. *Problems, Resources, and Issues in Mathematics Undergraduate Studies , 23* (2), 173-192.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impace of undergraduate science course innovations on learning. *Science , 331* (6022), 1269-1270.

Savin-Baden, M., & Major, C. H. (2004). *Foundation of problem-based learning.* Maidenhead, UK: Open University Press.

Schumacher, C. E. (2010, June 29). *Instructor's resource manual for use with Chapter Zero - Fundamental Notions of Abstract Mathemeatics, 2e.* Retrieved August 7, 2013, from http://www2.kenyon.edu/Depts/Math/schumacherc/public_html/Professional/Research/Zero/guide.pdf

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research , 69* (1), 21-51.

Yoshinobu, S., & Jones, M. (2013). An overview of inquiry-based learning in mathematics. *Wiley Encyclopedia of Operations Research and Management Science* , 1-11.

CONSIDERING MATHEMATICAL PRACTICES IN ENGINEERING CONTEXTS FOCUSING ON SIGNAL ANALYSIS

Reinhard Hochmuth[1], Rolf Biehler[2], Stephan Schreiber[1]
[1]Leuphana University of Lüneburg, [2]University of Paderborn

In the light of a rough description of the different contexts in which mathematics is learned and used in engineering studies, this report addresses epistemic relations between mathematics in higher mathematics lectures and mathematics in advanced engineering courses. In particular it elaborates on how different meanings of symbols, as subjectively relevant aspects of mathematical objects, are related to different institutional contexts and their dominant discourses. It is argued that modeling cycles are not an adequate tool in this context. Instead, we suggest using concepts from Anthropological Theory of Didactics (ATD). Inspired by (Castela & Romo Vázquez, 2011), exemplarily concepts from ATD are applied to topics and data from signal analysis. Finally, we claim this research could serve as a step towards investigating empirical questions relevant to students' learning and competences and, in particular, optimizing curricula and teaching in undergraduate mathematics.

Key words: Engineering, Higher Mathematics, Epistemological Aspects, Modeling Cycles, Anthropological Theory of Didactics

Introduction

It is well-known that mathematics plays an important role in engineering studies. All students have to pass basic courses in higher mathematics, and in theoretically-oriented engineering courses, mathematics is one of the major tools and obstacles. There is a general consensus that these difficulties with mathematics are a major reason for the high dropout rate in engineering studies, and thus, the teaching and learning of mathematics in engineering studies should be improved.

The German project KoM@ING[1] addresses this problematic situation by exploring the following question: Which topics, concepts, heuristic strategies and competences are relevant for being successful in basic theoretical courses like "Technical Mechanics" or "Theoretical Basics in Electrical Engineering" and in more advanced courses like "Systems and Signals" or specific lab courses? Answers to these questions are worked out by a specific combination of quantitative IRT-models and more qualitative process-oriented studies. The latter perspective causes a need for developing approaches that allow us to analyze students' use of mathematics in engineering contexts.

In the light of a rough description of the different contexts in which mathematics is learned and used in engineering studies, this report focuses in particular on epistemic relations between mathematics in higher mathematics courses and mathematics in engineering courses and how these relationships are reflected in two well-established approaches: modeling cycles and the Anthropological Theory of Didactics (ATD) (Chevallard, 1992, 1999). We focus on both of these approaches for the following reasons: Nowadays modeling cycles are very prominent and claim to conceptualize an important competence regarding school mathematics and, in particular, the application of mathematics to the "real world" and the role of mathematics in our daily lives. Assuming that students develop modeling competences in school, we question whether the use of mathematics in engineering studies could also be grasped by this approach, particularly in the transition from school to university. We show that, at least in our context regarding epistemological relations,

[1] This research was supported by BMBF 01PK11021D.

modeling cycles are generally not an adequate tool.[2] We then discuss concepts from ATD, since it claims to allow us to reflect on the use of mathematics within different institutional contexts, conceptualize related epistemological issues, and sketch ATD's potential regarding two different examples from signal analysis. Finally, we summarize our main conceptual results and propose possible subsequent research.

Mathematics in Engineering Studies

Students in engineering courses learn mathematics in at least three contexts. First, they have to pass courses in higher mathematics. Here the students learn mathematical concepts from analysis, linear algebra and sometimes elementary numerical analysis. These topics are mostly presented in a more or less theoretical mathematical setting. Rarely concrete applications relevant to engineering are presented.

Second, students must apply mathematics in their basic engineering courses. Since most of the mathematical concepts required in these courses have not been presented in the higher mathematical courses until that moment, often additional seminars are offered that accompany the engineering courses. But in those seminars the presentation is as a rule not as mathematically precise as in the mathematics courses.

While exact mathematical definitions and/or justifications for the mathematical concepts used in the basic theory-oriented engineering courses are often presented later in the higher mathematics courses, this is generally not the case for the more advanced mathematical concepts applied in courses like signal analysis. For example, a concept like delta-distribution is typically not covered in the mathematical courses attended by an electrical engineering student.

It is not clear how, if at all, students are able to integrate these variations of mathematics. To study this problem, it would be helpful to have methods that represent, relate and reflect these variations under a didactical perspective. The most prominent school-related approach that claims to conceptualize "mathematics in applications" is based on modeling cycles.

Modeling Cycles

All modeling cycles start with a situation where a problem or task from the world outside of mathematics has to be "solved". (For this paragraph see for example (Blum & Leiss, 2005; Haines & Crouch, 2010).) In a first step, or, depending on the concrete modeling cycle, in a series of steps, the problem has to be translated or transformed into a mathematical problem. This mathematical problem is then solved within the "mathematical world" by mathematical manipulations. Finally, the mathematical outcome has to be translated back to the "rest of the world" (that is, the world outside mathematics) and evaluated in view of the starting point. It is well-known that even in the case of modeling tasks in school, i.e. tasks which are specifically designed for applying some version of a modeling cycle, solution processes often do not follow such a cycle exactly. The stated purpose of this approach is that it permits the researcher to describe all possible, meaningful sequences of particular cognitive activities in solution processes (descriptive function) and it represents a helpful heuristic strategy for treating modeling tasks (normative function).

In our context, the crucial point is that, whether the tasks or situations in engineering studies are reasonable or not, students do not have to create the mathematical models by themselves. It is our impression that the tasks in signal analysis represent either mathematical problems or a problem that is already mathematically formulated in a signal analysis model. These models lie at the core of signal analysis as scientific discipline and are described, justified and explained in lectures and books. From the beginning, signal analysis and in

[2] Whether our arguments and conclusion interfere with the significance of this concept for school mathematics is besides the scope of this paper.

general also other topics in engineering courses are presented in a mathematically formulated and evenly formatted world that is the result of a scientific development that has happened through the course of history. The student has to act in view of this historically-evolved result but does not have to (re-)create it.[3]

Let us add some remarks in view of epistemological considerations worked out by (Wahsner & Borzeszkowski von, 1992) regarding mathematics in physics. They argue that, historically, specific physical achievements include figuring out which "behavior" can be considered as "property" and as "measurable" quantity, and, most important in our context, the physically adequate but often mathematical inconsistent use of mathematical symbols. Analogously, for an engineer or engineering student the symbols primarily represent context-related and context-embedded objects. For example, differentials are not understood as (in some sense) "ideal" infinitesimals or elements of a cotangent space but as finite, "physical" and even "measurable" quantities. We should expect students to treat these objects as a kind of physical quantities, since this allows students to do things with these objects that are as such mathematically inconsistent but enable them to understand and solve engineering tasks (see also our later discussion of the delta distribution).

Thus, there is not only no point "in the rest of the world", but there is also no point in a pure mathematical world. Therefore, modeling cycles seem to be not an adequate tool for grasping either the engineering or the mathematical side of tasks and solution processes.[4]

Assuming that the successful engineering student has to learn such an inconsistent use of symbols, an important but open question is how, and if, students and experts experience precisely this somehow contradictory (or better "dialectic") relation between different uses and meanings of mathematical symbols.

The very specific different meanings are realized or constituted within discourses related to the aforementioned different institutional contexts. The mathematical meaning is addressed in higher mathematics, and the context-specific meaning is mostly addressed in engineering courses such as signal analysis. An approach that claims to allow reflecting mathematics in different institutional contexts is ATD. Therefore we see potentialities in using ATD for reinterpreting modeling cycles in a more adequate way. In the next section, we will concentrate on ATD for better understanding the mathematical practices in signal analysis.

ATD and Signal Analysis

ATD (Chevallard, 1992, 1999) aims at a precise description of knowledge and its epistemic constitution. In the background of this approach is the conviction that a cognitive-oriented approach tends to misinterpret contextual or "institutional" aspects as personal dispositions, which sometimes obstructs the view on possible interventions.

A basic concept of ATD are praxeologies, which are represented in so called "4T-models (T,τ,θ,Θ)" consisting of a practical and a theoretical block. The practical block (know how, „doing math") includes the type of task (T) and the relevant solving techniques (τ). The theoretical block (knowledge block, discourse necessary for interpreting and justifying the practical block, „spoken surround") covers the technology (θ) explaining and justifying the used technique and the theory justifying the underlying technology (Θ).

In the end ATD ends up with local and regional mathematical organizations that allow contrasting and integrating practical and epistemological aspects of mathematical objects in view of different "institutional" contexts. We expect that this approach is in particular helpful in analyzing mathematical knowledge and its transformation to and within the three learning contexts of students in engineering courses mentioned above. This expectation is supported

[3] This argument as such does not necessarily contradict a constructivist perspective on learning.
[4] The importance of modeling cycles in school mathematics, its related "discourse" and what it reflects and constitutes, goes beyond the scope of this report.

by related but differently focusing ATD analyses in (Castela & Romo Vázquez, 2011) considering teaching signal analysis topics in mathematics and two control theory courses.

Next we sketch exemplarily an ATD analysis of the introduction of the delta-distribution in (Girod, Rabenstein, & Stenger, 2007)[5]: In a specific linear and time-invariant system it is asked, whether there an input signal can be found that produces the "system function" as output signal (T). It turns out that one could introduce a sequence of rectangle impulses, which produces a sequence of output signals converging to the "system function" of the specific system. Whereas the convergence of the output signals could be understood in a pointwise limit sense, hence in a sense coherent with a concept covered by higher mathematics courses, the "limit" of the rectangle input signals is a new object, which could mathematically be understood as a functional operating on test functions. But this understanding is generally beyond the scope of an introductory signal analysis lecture. Instead it is written "$\lim_{T_0 \to 0} x_{T_0}(t) = \delta(t)$", where the Dirac impulse δ is addressed as the pointwise limit function $\delta(t) = \begin{cases} 0 & \text{für } t \neq 0 \\ \infty & \text{für } t = 0 \end{cases}$. Additionally the effect of the Dirac impulse is symbolically represented by the integral $\int_{-\infty}^{\infty} f(t)\delta(t)dt = f(0)$, which relates to the concept image of an integral as infinite sum of infinitesimal small pieces $f(t)dt$ weighted by the "function" $\delta(t)$. In terms of ATD, applying the sequence of rectangle impulses etc. could be seen as τ (technique), "justifying" the limit of the infinitely narrow and infinitely high rectangle impulses via aspects of the concept image of the "limit"- and the "integration"-discourses represent θ (technology, discourse). Facets of Θ (theory) are graphical visualizations that are connected with the effect idea of integration and in particular symbolically based "analogies" to former learned mathematics.

The crucial point is that the signal analysis technique τ does not fit with higher mathematical discourses (technologies). The students have somehow to "learn" that they should neglect specific aspects from those discourses, for example they have to ignore concept definitions (The Riemannian integration of the Dirac impulse is not possible etc.) and at the same time, they have to realize aspects from them, for example specific parts of the "concept image" of integration and limits.

Finally we shortly discuss preliminary results from our qualitative study in KoM@ING: Based on teaching materials from several signal analysis courses we analyzed solutions to tasks which were given as voluntary homework in a signal analysis course at the University of Kassel[6]. A typical "simple" task looks at follows:

Let be given a low-bass-bounded signal $s(t)$ *with Fourier transform* $F\{s(t)\} = \begin{cases} 1, & \text{for } |f| \leq W \\ 0, & \text{for } |f| > W \end{cases}$. *Considering* $s(t)$ *as input signal classify the following assertions as true or false and justify your answer: i) The band-with of the output signal generated by a linear system is always less or equal to W. ii) The band-with of the output signal generated by a linear time-invariant system is always less or equal to W.*

Ad i) Solution based on mathematical techniques $(\tau_{sig,math})$ from the course: The system function $h(t) = \cos 2\pi f_0 t = \frac{e^{i2\pi f_0 t} + e^{-i2\pi f_0 t}}{2}$ gives for $s(t)$ by $F\left\{e^{i2\pi f_0 t}\right\}(f) = \delta(f - f_0)$ and

[5] For another analysis of this topic we refer to (Castela & Romo Vázquez, 2011, pp. 114-116).

[6] The authors are grateful to Prof. Dahlhaus (University Kassel) for placing the tasks and student solutions at our disposal.

$$F\left\{e^{i2\pi f_0 t}\cdot s(t)\right\}(f)=F\left\{e^{i2\pi f_0 t}\right\}*F\left\{s(t)\right\}(f)=\int_{-\infty}^{\infty}\delta(\tau-f_0)\chi_{[-W,W]}(f-\tau)d\tau=\chi_{[-W,W]}(f-f_0),$$

the output signal with spectrum $\frac{1}{2}\left\{\chi_{[-W,W]}(f-f_0)+\chi_{[-W,W]}(f+f_0)\right\}$, which represents a signal with a shifted band-with. Hence the assertion i) is false. An answer using a more intuitive signal analysis argument $(\tau_{sig,sig})$ would be: In a general linear system the transfer function changes in time, which could induce new frequencies and change the band-with.

Ad ii) "$(\tau_{sig,math})$-type solution": For a linear and time-invariant system the output signal of the input signal $g(t)$ is given by $F\left\{L(g)(t)\right\}(f)=H(f)F\left\{g(t)\right\}(f)$.This relation implies that the assertion is true, since $\operatorname{supp} F\left\{L(g)(t)\right\}\subseteq \operatorname{supp} F\left\{g(t)\right\}$. "$(\tau_{sig,sig})$-type solution": Since the transfer function does not change in time, no new frequencies arise.

In view of the approaches presented in the course lectures and the course material we expected in part ii) that both types of solutions arise in the students' homework. In fact mainly solutions that fit into the mathematical-type were given. In the subsequent observational study also $(\tau_{sig,sig})$-type solutions were presented. It's interesting that the students mostly supplemented a $(\tau_{sig,math})$-type solution.

All reasoning trials to part i) are of mathematical type. While the students' homework solutions are far from being correct and complete, the answers in the observational study (shortly after the exam) were mainly correct. Despite the complexity of the $(\tau_{sig,math})$-type solution no student tries a $(\tau_{sig,sig})$-type solution.

Both examples illustrate that ATD might allow identifying aspects and major differences in concepts and students' solution processes that are strongly related to epistemological issues and institutional contexts. We expect that the ongoing data collection (in particular problem focused guided interviews) will shed further light on our preliminary observations leading to a deeper understanding of students' mathematical practices in engineering contexts.

Summary and outlook

We discussed certain aspects of the use of mathematics in engineering studies. Our question was how are different mathematical practices in higher mathematics lectures and in advanced engineering lectures grasped. This relates to the different meanings of symbols and how they are related to different institutional contexts and their dominant discourses. We argued that modeling cycles are not an adequate tool in this context. Instead, we suggested using concepts from ATD and applied them to data from signal analysis.

It can be expected that, in solving a specific task, students have to make specific decisions regarding the relevance of knowledge, in terms of ATD "technic, technology and theory". In (Tuminaro & Redish, 2007) task-specific "decision" processes were studied from a cognitive point of view. Combining these ideas with ATD and taking into account culture-historical approaches like "communities of practices" (see for example (Lave, 1988; Wenger, 1998)) could shed further light on those subjectively localized but essentially societally mediated "decision" processes. We believe that those processes include in particular sub-jectification mechanism (see for example (Brown, 2008a)) that are constituted by discourses, which are itself (dialectically) related to mathematical praxeologies.

Following this line, ATD analyses could serve as a step towards investigating empirical questions such as: Which meanings are actualized in the context of specific tasks and situations? In the end, answers to those questions will give hints for optimizing curricula and teaching of undergraduate and higher mathematics courses as well as signal analysis.

References

Blum, W., & Leiss, D. (2005). How do students and teachers deal with mathematical modelling problems? The example "Sugarloaf". *ICTMA 12 Proceedings* (pp. 222-231).

Brown, T. (2008a). Lacan, subjectivity and the task of mathematics education research. *Educational Studies in Mathematics, 68*, 227-245.

Castela, C., & Romo Vázquez, A. (2011). Des Mathematiques a l'Automatique: Etude des Effets de Transposition sur la Transformee de Laplace dans la Formation des Ingenieurs. . *RDM, 31*(1), 79-130.

Chevallard, Y. (1992). Fundamental concepts in didactics: Perspectives provided by an anthropological approach. *RDM, Selected Papers. La Pensée Sauvage, Grenoble*, 131-167.

Chevallard, Y. (1999). L'analyse des practiques enseignantes en théorie anthropologique du didactique. *RDM 19*(2), 221-266.

Girod, B., Rabenstein, R., & Stenger, A. K. E. (2007). *Einführung in die Systemtheorie - Signale und Syteme in der Elektrotechnik und Informationstechnik.* Wiesbaden: B.G. Teubner Verlag.

Haines, C. R., & Crouch, R. (2010). Remarks on a Modeling Cycle and Interpreting Behaviours. In R. Lesh et al. (Ed.), *Modeling Students' Mathematical Modeling Competencies* (pp. 145-154): Springer.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life.* Cambridge: Cambridge University Press.

Tuminaro, J., & Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physics Education Research, 3*, 1-22.

Wahsner, R., & Borzeszkowski von, H.-H. (1992). *Die Wirklichkeit der Physik. Studien zur Idealität und Realiät in einer messenden Wissenschaft*. Frankfurt/M., Berlin, Berlin: Peter Lang Vlg.

Wenger, E. (1998). *Communities of practice*. Cambridge: Cambridge University Press.

DIFFERENCES IN EXPECTATIONS BETWEEN EXPLICIT STATEMENTS AND ACTUAL PRACTICES USING VECTORS IN A TRIGONOMETRY AND PHYSICS COURSE

Wendy James
University of Central Oklahoma

Science and engineering instructors often observe that students have difficulty using or applying prerequisite mathematics knowledge in their courses. Historically, transfer theory is used to investigate students' issue applying their vector knowledge from a trigonometry course to a physics course, but this qualitative case-study is positioned differently epistemologically and theoretically from transfer theory to understand and describe the mathematical vector practices in the two courses. Saussure's (1959) concept of signifier and signified provided a lens for examining the data during analysis. Multiple recursions of within-case comparisons and across-case comparison were analyzed for differences in what the instructors and textbooks explicitly stated and later performed as their practices. While the trigonometry and physics instruction differed slightly, the two main differences occurred in the nature and use of vectors in the physics course.

Key words: Vectors, Trigonometry, Physics, Semiotics, Transfer, Literacy

Background of the Problem

Science and engineering instructors often complain that the prerequisite math courses do not prepare the students for their courses, and as a result, the instructors feel they still have to teach the mathematics along with the science and engineering. Knight (1995) offered a study in which 86% of the students in his study reported remembering that they had studied vectors prior to the physics course, but when their knowledge was evaluated, only a third of the students came with sufficient knowledge, and "a full 50% entered with no useful knowledge of vectors at all" (p. 77). In addition, Nguyen and Meltzer (2003) found that even after a full semester of physics "more than one quarter of students beginning their second semester of study in the calculus-based physics course, and more than half of those beginning the second semester algebra-based sequence, were unable to carry out two-dimensional vector addition," (p. 630). While these studies contribute evidence that students lack the requisite mathematical vector knowledge, they do not explain why the phenomenon occurs. Further research is necessary to seek an explanation, which this project contributes.

Historically transfer theory was used to explore why students have a difficult time applying their prerequisite knowledge to a new context. The problem with traditional transfer theory is that it assumes that students learned what they were supposed to learn with varying degrees of acquisition, but over the past decades, researchers have gradually adjusted their assumption about student learning from students *acquiring* knowledge to students *constructing* knowledge (Kieran, Forman, & Sfard, 2001). The change in assumptions about how students learn is accompanied by changes in assumptions about the nature of mathematics itself. Historically, mathematical symbols have been viewed as having fixed referents with the capability of embodying those fixed referents, but more recently math researchers recognize the multiple, nuanced meanings symbols hold depending on the context in which they are used (e.g. Sfard, 2000). Learners construct the meaning of a symbol much like the meaning of any word: through context and use (Sfard, 2003).

Sometimes instruction provides intentional and unintentional messages about the meaning of symbols. For example, "The symbol which is used to show equivalence, the equal sign, is not always interpreted in terms of equivalence by the learner" (Kieran 1981, p. 317). Because students repeatedly see the equal sign separating the problem from the answer and representing the operating button on a calculator, they come to believe the meaning of an equal sign is operational—a "do something" signal. Jones and Pratt (2005), Falkner, Levi, and Carpenter (1999), Saenz-Ludlow and Walgamuth (1998) found that in adjusting the student-learning activities, students in their studies seemed to adopt equivalence as the meaning of an equal sign. The mathematics class activities were unintentionally causing students to interpret the meaning of the equal sign differently than would have been explicitly stated by the instructor.

Students use the symbols and symbol systems even before they know exactly what the symbols mean and signify, and as a result, the student also constructs the meaning of the symbols from the process of using them (Sfard, 2000, 2003). Borrowing ideas expressed in literacy theories, students construct their understanding of different mathematical literacies, and these literacies may vary across communities of practice. Wenger (2006) writes, "Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly." Because separate communities of practice—even within the same specialized area—exist, separate literacies form, and over time, these communities of practice adjust their content, processes, and ways of thinking resulting in the literacies being historically-contingent and evolving.

In a mathematics class, instructors are seasoned members of the mathematical community of practice, and students are required to learn both the mathematical content and processes, and the accompanying symbols and manners of symbolizing required by the content and processes. Likewise, in a physics class, instructors are seasoned members of the physics community of practice, and students are required to learn both the mathematical content and processes and the physics content and processes and the accompanying symbols and manners of symbolizing. The aspects of what instructors know and students must learn are the very elements necessary to be considered mathematically literate in the particular subject matter being taught. This paper contributes a description of some of the practices concerning vectors used during instruction for a trigonometry and a physics course in order to offer the reader ideas concerning what may contribute to students' difficulties with vectors.

Methodology

The purpose of this project was to begin to characterize the various practices of two academic disciplines, specifically trigonometry and physics, with respect to the concept of vectors and to describe any differences in their practices. These practices are modeled and described in course instruction; therefore, the research design for this project required accessing and objectifying the instruction for analysis while not stripping the instruction of its complexity (Lemke, 1998; Patton, 2002).

Videos from one college trigonometry course and one college, algebra-based physics course were selected for analysis, and the two courses were designed as separate case studies for in-depth study and comparison (Patton, 2002). The trigonometry course had a two-day vector unit; both days were videoed. Physics uses vectors throughout the course; this project videoed instruction up through the first exam, which covered the first three chapters of the textbook (mathematical chapter, a 1-D motion chapter, and a 2-D motion chapter). The trigonometry instructor had an engineering background, and the physics instructor had a mathematical background. Instructor interviews, live observations during the course lecture taping, and course textbooks served as qualitative data to support and complicate the analysis and patterns stemming from course videos.

Borrowing language and ideas from Saussure (1959), mathematical symbols, objects, and vocabulary could all be considered to be signs. Each sign has two parts: its signifier and its signified. The signifier is the visually accessible form of the sign, and the signified is the concept and/or meaning that is being represented by the signifier. For example, "½" or "half" are both signifiers. The use of the numbers 1 and 2, where 1 is above the 2 and has a line between them, signifies the same thing as the word using the letters *h,a,l,f*, and in both cases, what they signify is the quantity of half of an object or half of a set of objects. Saussure's (1959) concept of the duality of a sign in having both a signifier and signified was used as a way of examining the data during analysis. Signifiers for vectors (notations, graphs, algebraic expressions, …) were collected and compared across instruction. The signifiers were then analyzed for patterns in use for intended signification.

Multiple recursions of within-case comparisons and across-case comparison were analyzed for differences in what the instructors and textbooks explicitly stated and later performed as their practices. All signifiers introduced by the instructors concerning vectors (notation, diagrams, and vocabulary) were individually analyzed for patterns in their meaning and use within and across the instruction.

Analysis

When the physics instructor and textbook teach a mathematical lesson on vectors prior to introducing the physics unit, their lessons are very similar to the instruction provided by the trigonometry instructor and textbook. Both the trigonometry and physics instructors spent part of two lectures introducing what vectors are, styles of notation, how to express vectors algebraically and graphically, and how to perform algebraic and graphical operations with two-dimensional vectors. The instruction was very similar, but not all together identical. For the brevity of this paper, a description of their similarities and differences are not fully described here. To summarize, the instruction had strong agreement in introducing vectors as geometric objects that would be sketched as arrows, that vectors could be labeled and named with notation such as $\vec{A}$, that the purpose of using vectors is to represent paired information, and that vectors could be added, subtracted, and multiplied by a scalar. Despite the similarities across the two courses' mathematical instruction toward vectors, three interesting conflicts between explicit statements and the actual practices while "doing physics" surface: vectors are not always graphed as arrows, vectors are not always expressed algebraically and graphically as paired information, and vectors do not always extend "from one place to another in the coordinate plane." The following paragraphs provide a brief explanation of each.

When introducing vectors, instructors' explicit statements introduce vectors as geometric objects depicted as arrows. The trigonometry instructor states vectors are defined as "directed *line segments*" (emphasis by researcher) as she draws line segments and adds arrows to indicate their directions. The physics instructor began his unit wanting his students to recognize "that there are two major kinds of structures" that are used in physics: scalars and vectors. He explained how most students do not see vectors and scalars as different from one another; therefore, his instruction began by describing scalars as quantities and vectors as quantities requiring the additional description of direction within a referenced system. He explained mass, temperature, and weight can be measured as a singular quantity, which he calls scalars, but position, velocity, and acceleration would require them to use vectors. He then transitioned to begin his instruction on the mathematics of vectors. Similar to the trigonometry instructor, the physics instructor referenced a diagram of an arrow while he defined vectors. In two informal verbal and one written statements, the physics instructor states “vectors go/point/extend from one point/place to another/other in the coordinate system/plane.”

Despite these explicit statements, once the physics unit begins, vector quantities are depicted for almost all of the next two lectures without being depicted in graphs as arrows. For example, in Figure 1 velocity is not graphed as an arrow in the velocity-verses-time graph. Instead, velocity is being depicted simply as various numerical values. Further research is needed to explore whether students’ expectations from explicit statements during instruction that vector quantities should be graphed as arrows cause difficulties negotiating the meaning of graphs having vector quantities as numerical values paired with time.

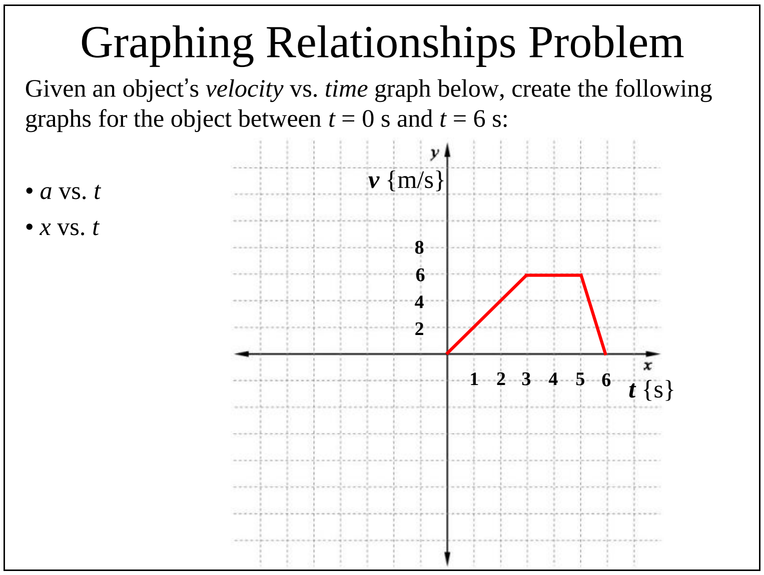


Figure 1. Excerpt from the physics instructor’s lecture. PowerPoint slide 5 on day 3 of analysis does not depict vector quantities as arrows—rather they are numerical values paired with time in the velocity-verse-time graph.

A second difference between what is explicitly stated by the instructors and then later performed when “doing physics” is the manner in which vectors are algebraically written. When the trigonometry and physics instructor introduce vectors, they both introduce several styles of notation. Some notation is used as a means of naming vectors, and some notation is used as a means of algebraically describing vectors. The trigonometry instructor introduces $\overrightarrow{PQ}$ and $\vec{v}$ as ways of naming vectors and paired coordinates P(-2, 4)Q(4, 7), rectangular coordinates $\langle 6,3 \rangle$, and $6\vec{\imath} + 3\vec{\jmath}$ as ways of algebraically describing vectors. Similarly, the physics instructor introduced $\vec{v}$ as a way of naming vectors and rectangular coordinates (x, y), polar coordinates (R, θ), and $A\hat{x} + B\hat{y}$ as ways of algebraically describing them. Both instructors emphasize that vectors are described algebraically using paired information and can be written with one of these manners of notation.

Despite explicit statements introducing vectors as paired information and the various styles of algebraic expressions, the physics instructor and textbook do not use these styles of notation to express vectors algebraically in the Kinematic unit. Instead, two-dimensional vectors are broken into their components and labeled using component notation, and one-dimensional vectors are simply labeled with component notation. Throughout the trigonometry instruction, vectors are always algebraically depicted using one of the given notations.

The graphs in Figure 1 and Figure 2 provide two samples of the physics instructor's manner of referencing vector quantities as scalars. Figure 1 provides an example of a velocity-verses-time graph in which velocity is depicted as scalar values. In addition, Figure 2 has the beginning of the physics instructor's board work while solving a two-dimensional kinematic problem. Notice the vectors in the sketch are being labeled with component notation, and the vectors have been listed in the table to the right as separate components—not using the algebraic notation taught in the previous math-centered lessons. For example, acceleration is not written using the previously introduced style of algebraic notation, such as $\vec{\boldsymbol{A}} = 0\hat{x} - g\hat{y}$ or $\vec{\boldsymbol{A}} = -g\hat{y}$; instead, the table describes the acceleration as two independent variables as $A_x = 0$ and $A_y = -g$. The majority of all the problems by the physics textbook and instructor referenced acceleration just as $a_y = -g$ with no reference to the x-component at all. As the physics instructor continues to solve this problem, Kinematic equations are used, and these numerical values in the table are substituted into the equations. During none of the process of working Kinematic problems are vectors algebraically written as paired information nor are vector operations performed.

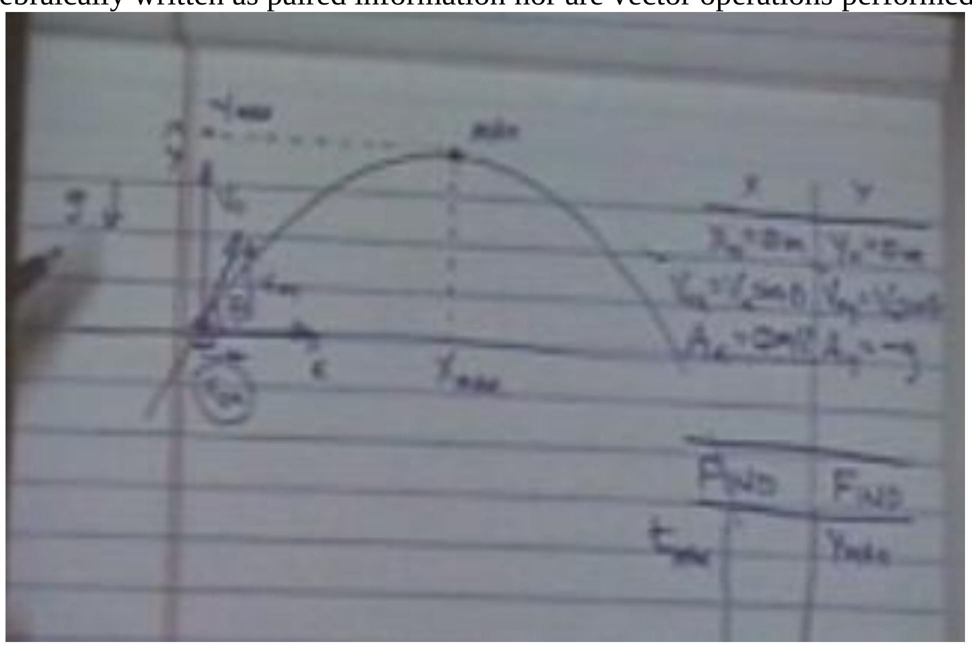


Figure 2. Excerpt from the physics instructor's lecture. Setting up to solve a two-dimensional projectile motion problem using Kinematic equations. Notice arrows are not labeled with vector notation, and the table describes vectors as two independent variables (components) rather than using vector notation.

A third difference between what is explicitly stated by the instructors and then later performed when "doing physics" concerns a quality of the vectors. When introducing vectors, the trigonometry instructor, the trigonometry textbook, the physics instructor, and physics textbook all seem consistent in conveying that vectors are geometric objects that *begin and end at particular places*. The trigonometry instructor states vectors are defined as "directed *line* *segments*" (emphasis by researcher), and she clarifies that line segments have *end points* (line 26), which is where the vectors begin and end. The physics instructor states in his two verbal and one written statements that vectors go/point/extend from *one point/place* to *another/other* in the coordinate system/plane (emphasis by researcher). Similarly, the physics textbook states "The vector $\vec{A}'$ *from point 3 to point 4* … has the same length and direction as the vector $\vec{A}$ *from point 1 to point 2*" (italics added by researcher).

Despite statements introducing vectors as beginning and ending at particular points/places, the manner in which vectors are *used* to represent some vector quantities seems to contradict the explicit statements. Vectors that represent force, velocity, or acceleration do not go from one point to another in the coordinate plane; whereas, vectors that represent displacement or that are context-free go "from one endpoint to the other" (see Figure 3).

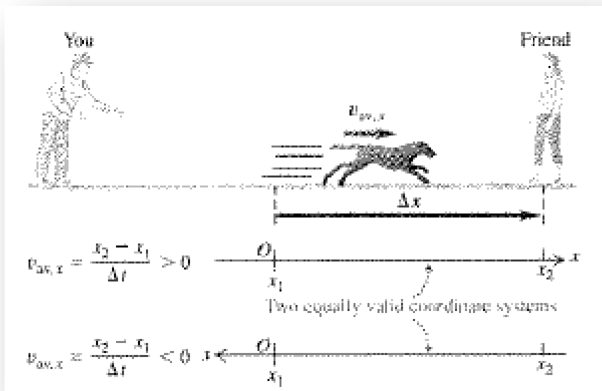


Figure 3. Excerpt from p. 33 of physics textbook offering an example of a displacement vector extending from x_1 to x_2 while the velocity vector is not fixed to a particular location on the coordinate axes like the displacement vector is.

Because the physics instructor and textbook use displacement vectors and context-free vectors while they introduce vectors, their original statements accurately describe the original vectors, but later when vectors represent velocity and acceleration, vectors no longer "extend from one place to another in this coordinate plane."

Discussion.

Using literacy as a metaphor for students participating in communities of practice to develop the ways of speaking, reading, and doing mathematics seemed to elucidate some differences that may have gone unseen by earlier research. Mathematics is not a fixed concept and, as such, can develop multiple meanings and uses in separate communities of practice. This project assumed the possibility that the mathematics of vectors used by a trigonometry course and a physics course may have differences that effect student learning success. The differences between what was explicitly stated and then practiced was not a result of poor teaching. The textbooks, which are created by multiple authors with innumerable revisions followed the same patterns as the instructors. These differences seem to be insignificant to people who are seasoned members of the community, but they may be difficulties for new learners.

Sfard (2003) writes, "The act of naming and symbolizing is, in a sense, the act of inception, and using the words and symbols is the activity of constructing meaning" (p. 374).

When students observe the activities using vectors, they observe different practices of reading and writing vectors as compared to the practices explicitly stated in strictly mathematical situations. The initial moments of inception in which arrows and notation are infused with meaning differ from the meanings the arrows take later and the manner in which the notation is later used. Is it possible that separate mathematical literacies between trigonometry and physics communities is causing students problems? Can conversations between instructors help bridge the gap in differences in the manner in which the mathematics is used? This study only observed vector use during the first two chapters of instruction when motion along a straight line and in a plane was being studied. Are there other differences in the use of the mathematics that might further strengthen an argument toward separate mathematical literacies? Do these differences impact students as they learn physics? Further conversations and research is encouraged.

References

Falkner, K.P., Levi, L., & Carpenter, T.P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics, 6*(4), 232.

Jones, I. & Pratt, D. (2005). Three utilities for the equal sign. In Chick, H.L. &Vincent, J.L. (Eds.) *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education,* (Vol. 3, pp. 185-192). Melbourne: PME.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*, 317-326.

Kieran, C., Forman, E., & Sfard, A. (2001). Learning discourse: Sociocultural approaches to research in mathematics education. *Educational Studies in Mathematics, 46*, 1-12.

Knight, R.D. (1995, Feb). The vector knowledge of beginning physics students. *The Physics Teacher, 33*, 74-77.

Lemke, J.L. (1998). Analysing verbal data: Principles, methods, and problems. In K. Tobin & B. Fraser (Eds.). *International Handbook of Science Education* (pp. 1175-1189). London: Kluwer Academic Publishers. Retrieved February 16, 2009 from http://academic.brooklyn.cuny.edu/education/jlemke/papers/handbook.htm

Nguyen, N. & Meltzer, D.E. (2003). Initial understanding of vector concepts among students in introductory physics courses. *American Journal of Physics, 71*(6), 630-638.

Patton, M.Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Saenz-Ludlow, A. & Walgamuth, C. (1998). Third graders' interpretations of equality and the equal symbol. *Educational Studies in Mathematics, 35*, 153-187.

Saussure, F. de. (1959). *Course in general linguistics*. (C. Bally, A. Sechehaye, with A. Riedlinger, Eds.; W. Baskin, Trans.). New York: McGraw-Hill.

Sfard, A. (2000). Symbolizing mathematical reality into being—or how mathematical discourse and mathematical objects create each other. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sfard, A. (2003). Balancing the unbalanceable: The NCTM Standards in light of theories of learning mathematics. In J. Kilpatrick, W.G. Martin, and D. Schifter (Eds.), *A research companion to principles and standars for school mathematics* (353-392). Reston, VA: NCTM.

Wenger, E. (2006). Communities of practice: A brief introduction. Retrieved from http://www.ewenger.com/theory

Proof Conceptions of College Calculus Students

Introduction

There has been “increasing awareness that reasoning is central to mathematics and mathematics learning” (Yackel & Hanna, 2003, p. 227) among education researchers, which raises questions about student conceptions of reasoning, argumentation and mathematical proof. Mathematicians and mathematics education researchers have consistently asserted the crucial roles deductive reasoning and proof play in discovering, communicating, verifying, understanding, and systematizing mathematics (Hanna, 2000; Ko, 2008; Thurston, 1998). In response to the many affirmations of the importance of proof for learning and understanding mathematics, there has been extensive study of students’ conceptions of mathematical proof, their abilities to construct and understand proofs, and the frequency with which inductive evidence is accepted as sufficient verification of mathematical conjectures (Bell, 1976; Healy & Hoyles, 2000; Ko, 2008; Stylianides, 2009; Varghese, 2009).

These studies have provided valuable insights into “what types of reasoning students are capable of at various age and grade levels, how their notions of reasoning and proof develop over time, and what limitations in reasoning they exhibit” (Yackel & Hanna, 2003, p. 230). Researchers have focused on different populations, ranging from elementary school students to advanced university mathematics students, and they have yielded several consistent and useful findings. Broadly speaking, many students at all grade levels have difficulty with the processes of creating and evaluating deductive mathematical proofs, and many believe simple empirical arguments are convincing proofs.

Past research has focused primarily on three groups: students in high school geometry courses, pre-service and in-service teachers, and students in advanced undergraduate courses who have received formal instruction in proof writing. However, little attention has been given to examining students’ understanding of proof after the completion of a high school geometry course, but before taking a course explicitly focused and dependent upon proof. The purpose of this study is to fill this gap in the literature by examining university students enrolled in Calculus courses.

The following two research questions were used to frame the study:

1. What are college Calculus students' views and thinking about the nature and purpose of mathematical proof?

2. What forms of empirical arguments are accepted by Calculus students as proof of a mathematical conjecture?

The Common Core State Standards for Mathematics, or CCSSM, which have now been adopted by 45 states and three U.S. territories, emphasizes the importance of reasoning and proof in K-12 mathematics education. The CCSSM states that secondary school students must learn to “construct viable arguments and critique the reasoning of others,” to “reason abstractly and quantitatively,” and to begin “using more precise definitions and developing careful proofs” (National Governors Association, 2010, p. 74). The NCTM Standards (2000) make the more forceful claims that “systematic reasoning is a defining feature of mathematics” (p. 57), and that secondary students should “recognize reasoning and proof as fundamental aspects of mathematics” and “develop and evaluate mathematical arguments and proofs” (p. 342).

Studying students in an introductory college calculus courses can contribute to evaluating whether recent high school graduates have an understanding of mathematical reasoning and proof consistent with what is called for by the CCSSM and NCTM standards. By knowing which aspects of reasoning and proof remain particularly for this population, researchers and practitioners can more effectively develop and evaluate targeted interventions for K-12 mathematics students.

Methods

All data were collected and analyzed within a cognitivist theoretical framework, and it is therefore assumed that inferences can be made about the conceptual understandings students possess through the analysis of their voluntary responses to written and verbal prompts. Many previous investigations related to student understanding of mathematical proof have similarly used this framework in investigating student understanding of reasoning and proof with data from written questionnaires, in-person interviews, and classroom case studies. Thus, especially given that some of the analysis involves comparisons to findings from these prior studies, it is believed to be a suitable framework for the present study.

An eight-item written survey, titled the Mathematical Reasoning Questionnaire, comprised of mathematical tasks used during past investigations of the conceptions of mathematical proof of secondary school and undergraduate students was the primary data collection instrument. In the present study, only results from the first three survey items are presented and discussed. Coding schemas devised during prior studies of student proof conceptions were used to code student responses to questionnaire and interview responses. Adhering to the coding categories that have been used in the past when analyzing survey and interview data will permit comparisons to prior studies, and this allows for a discussion the significance of results in the context of previous work.

Participants were solicited to complete the questionnaire from an introductory differential Calculus course at mid-sized northeastern university. At this level in the typical undergraduate mathematics sequence, students are assumed to not have received substantial formal instruction in the creation and evaluation of mathematical proofs beyond what they may have received in secondary school. Further, for many students, Calculus may be their last purely mathematical formal educational experience, and thus their conceptions of mathematical proof may reach their peak development at this level. Fifty-two students completed the questionnaire. Four versions of the survey with differing question orders were used, but question order was not found to be a significant factor in responses.

Results and Analysis

The first item on the questionnaire, shown in Figure 1, has not been used in prior studies, but similar questions are common in introductory abstract mathematics courses. This item is designed to provide information about which strategies and forms of argument are favored by students during the construction of an elementary proof. The concepts of integer and parity involved are basic enough that all participants can be reasonably expected to be familiar with the terminology and underlying concepts. Responses were classified as using either inductive or deductive proof schemes, and were further differentiated by the criterion outlined in Table 1, which are modeled after a framework for classifying student proofs proposed by Bell (1976). Several categories pertaining to the enumeration and testing of all possible cases were removed from Bell's (1976) framework because they were not applicable to tasks involving infinite domain sets.

(1) Determine if the following statement is true or false and justify your answer: if x is odd integer, then 9x+2 is an odd integer.

Figure 1. Survey question one

Proof Scheme Category	Description	n	%
1: Systematic Inductive	Student believes several sets of cases must be tested and attempts to empirically test subsets from each set.	1	1.9
2: Non-systematic inductive with many examples	Student tests two or more discrete cases without explicit reason for testing the particular cases	18	34.6
3: Non-systematic inductive with one example	Student tests only one discrete case without explicit reason for testing that particular case.	9	17.3
4: Complete systematic and deductive	Student constructs a valid deductive proof using accepted axioms.	0	0.0
5: Systematic with gap and deductive	Student creates a logically valid argument, but appeals to principles that cannot be taken as true without further verification.	1	1.9
6: Partially systematic and deductive	Student analyzes the situation deductively and produces some relevant information, but fails to build a complete, connected argument.	8	15.4
7: Non-systematic and deductive	Student recognizes a need to verify deductively and attempts to represent the conjecture in general terms, but a subsequent argument is nonexistent or contains many serious errors.	4	7.7
8: Solves equation	Student solves the equation 9x+2 = 0 and attempts to draw a conclusion from the result.	4	7.7
No Response	Student chose not to answer the question or wrote only “true” or “false.”	7	13.5
		52	100

Table 1. Survey question one coding categories and responses

62.2% of respondents used inductive arguments and 37.8% used deductive arguments. While significant differences exist in the types of arguments used ($\chi^2 = 44.4, p < 0.01$), there is insufficient evidence to conclude that a statistically significant difference exists between the proportions of calculus students using inductive instead of deductive arguments (p > 0.1). However, it is clear that students are at least as likely to use purely empirical evidence as they are deduction, and the demonstration of several random cases, classified as proof scheme 2 in Table 1, was the most commonly form of argument relied upon for justification. Balacheff (1988) describes this as naïve empiricism, which “consists of asserting the truth of a result after verifying several cases” (p. 218).

The empiricist approach of testing many cases may be effective, although perhaps cumbersome, in some discrete systems where only a finite number of possibilities exist. However, many mathematical conjectures apply to infinite domains, and this makes naïve empiricism only a way to provide fragmented evidence in support of a conjecture rather than a method for producing a complete proof. This approach is inconsistent with the views of mathematicians and the demands of the CCSSM, and in situations where a counterexample is not immediately obvious, the naïve empiricist strategy can quickly lead students to make false conclusions.

Survey items two and three, which were first used in a study by Martin and Harel (1989) (1989), are based on the notion that for many students, the perceived validity of a mathematical argument is dependent upon superficial characteristics. Both are Likert-style items that require students to rate arguments for elementary integer divisibility theorems. Only item two, shown in

Figure 2, is shown due to space considerations. In both questions, the arguments to be rated are structured to follow the proof schemes outlined in Table 2. The purpose of including two items with such a high similarity is to gauge the level of consistency in student reasoning.

(2) Consider the following statement and arguments. Using the space provided, rate how well each argument convinces you of the truth of the statement on a scale of 1 to 4 using the following criteria:

1: Not a convincing mathematical proof
2: Slightly convincing mathematical proof
3: Mostly convincing mathematical proof
4: Completely convincing mathematical proof

Statement: If the sum of the digits of any whole number is divisible by 3, then the number itself is divisible by 3.

Arguments:
A. _______ The sum of the digits of 123 is 6, which is divisible by 3. The number 123 is also divisible by 3.

B. _______ We can pick any number so the sum of its digits is divisible by 3, say 731234082. The sum of the digits is 30, which is divisible by 3, and the number itself is also divisible by 3.

C._______ 31 is not divisible by 3, and the sum of its digits is 4, which is not divisible by 3. On the other hand, 36 is divisible by 3, and the sum of its digits is 9, which is divisible by 3.

D. _______ If we list several numbers that we know are divisible by 3, say 3, 6, 12, 15, 18, 24, 36, 48, 1002, 1008, and so on, we can see that the sums of their digits are always divisible by 3.

E. _______ Let a be any whole number such that the sum of its digits is divisible by 3. Assuming its digits are x, y, and z, then $a = xyz$. Since $x + y + z$ is divisible by 3, the xyz is also divisible by 3. Therefore, a is divisible by 3.

Figure 2. Survey question two

Proof Scheme	**Description**
Argument A: Single example	A single example supporting relationship is used as proof.
Argument B: Extreme example	An example of the relationship holding for case perceived as extreme, such as a large number, is used as proof.
Argument C: Example and non-example	The combination of evidence from a supporting example and a non-example are used as proof.
Argument D: Many examples	A list of many examples supporting the relationship is used as proof.
Argument E: False deductive	A false proof of the generalization. This is not truly a deductive proof, but may be viewed as such by students because of its ritualistic aspects (Vinner, 1983).

Table 2. Survey question two and three argument classifications

Consistent with the methods of Martin and Harel (1989), responses of "1" or "2" were categorized as not supporting the argument and responses of "3" or "4" were categorized as supporting the argument. The χ^2 goodness of fit test was used to compare the proportions in each category with the alterative hypothesis that the proportion supporting was greater than the proportion not supporting the argument. Results for items two and three are shown in Table 3 and Table 4, respectively. Survey question three did not include an argument based on an extreme case, so the argument B column is omitted from Table 4.

	A	B	C	D	E
Support argument	26	31	30	38	27
Do not support argument	24	19	20	12	23
χ^2, **p**	$\chi^2 = 0.08$, p>0.1	$\chi^2 = 2.88$, p<0.1*	$\chi^2 = 2.0$ p>0.1	$\chi^2 = 13.52$ p<0.01*	$\chi^2 = 0.32$ p>0.1

Table 3. Question two responses * Significant result

	A	C	D	E
Support argument	29	30	41	30
Do not support argument	22	21	10	21
χ^2, **p**	$\chi^2 = 0.96$, p>0.1	$\chi^2 = 1.59$ p>0.1	$\chi^2 = 18.8$ p<0.01*	$\chi^2 = 1.59$ p>0.1

Table 4. Question three responses * Significant result

The data indicate that most introductory calculus students will accept an inductive argument as a complete mathematical proof. A significant majority of students indicated that the testing of many cases is a convincing and complete form of mathematical proof. A significant majority also indicated that the testing of an extreme case, evidenced by argument B results in Table 3, is acceptable as proof. Martin and Harel (1989) similarly reported a significant acceptance of argument styles B and D in their study of 101 pre-service elementary teachers. These findings suggest that many students are not completing school school with a sufficiently robust view of the role of empirical evidence in mathematics, and that the treatment of these subjects in K-12 education and during college calculus must be examined.

Perhaps among the most concerning was the finding that the calculus students sampled were equally likely to accept false arguments that possessed superficial features of some deductive proofs, such as the use of variables and references to established mathematical principles, as they were to reject them. This phenomenon has been identified in past research by Vinner (1983) and Martin and Harel (1989), but little is known about which instructional practices can help students to form effective strategies for critically evaluating mathematical arguments in ways that expose nonsensical proofs. The data indicates that many calculus students evaluate and accepts proofs based on the form of the argument or the authority of its source rather than the coherence of the argument.

Proof construction and evaluation is clearly are clearly difficult skills to develop, so K-12 mathematics and university calculus instructors must be cautious to not leave students behind when dealing with these topics. A sizable proportion of students may report that they understand proofs used in class, which may lead instructors to incorrectly gage the proficiency of his or her students, when in reality the underlying mathematical arguments are largely being ignored.

Questions for the Audience

1. What other types of survey tasks could be used to provide insights into the research questions?
2. How could interviews be structured to learn more about the empirical proof schemes of calculus students?
3. What are the strengths and limitations of using college calculus courses as a proxy for studying the mathematical reasoning skills developed during K-12 education?

Works Cited

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. *Mathematics, teachers and children*, 216–235. Retrieved from http://edumat.uab.cat/Diseo/Balacheff.pdf

Ball, D., & Bass, H. (2003). Making mathematics reasonable in school. In J.Kilpatrick, W.G. Martin, and D. Schifter (Eds.), *A Research Companion to Principals and Standards for School Mathematics (pp.27-44).* Reston, Va; National Council of Teachers of Mathematics

Bell, A. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics*, *7*(1), 23–40. Retrieved from http://www.springerlink.com/index/G68V76300RL36K00.pdf

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational studies in mathematics*, 5–23. Retrieved from http://www.springerlink.com/index/Q21702846NV02615.pdf

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for research in mathematics education*, *31*(4), 396–428. Retrieved from http://www.jstor.org/stable/10.2307/749651

Ko, Y. (2010). Mathematics teachers' conceptions of proof: implications for educational research. *Journal of Science and Mathematics Education*, (August 2008), 1109–1130. Retrieved from http://www.springerlink.com/index/5155675230q21233.pdf

Martin, W., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education. Retrieved from* http://www.jstor.org/stable/10.2307/749097

National Governors Association Center for Best Practices, C. of C. S. S. (2010). Common Core State Standards for Mathematics. *Retrieved September*. Retrieved from http://www.corestandards.org/assets/CCSSI_Math Standards.pdf

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Schoenfeld, A. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for research in mathematics education*, *20*(4), 338–355. Retrieved from http://www.jstor.org/stable/10.2307/749440

Selden, A., & Selden, J. (2003). Validations of Proofs Considered as Texts: Can Undergraduates Tell Whether an Argument Proves a Theorem? *Journal for Research in Mathematics Education*, *34*(1), 4. doi:10.2307/30034698

Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics*, *72*(2), 237–253. doi:10.1007/s10649-009-9191-3

Varghese, T. (2009). Secondary-level Student Teachers' Conceptions of Mathematical Proof. *IUMPST: The Journal*, *1*(June), 1–14. Retrieved from http://www.k- 12prep.math.ttu.edu/journal/contentknowledge/varghese01/article.pdf

Vinner, S. (1983). The notion of proof–some aspects of students' views at the senior high level. In *Proceedings of the Seventh International Conference for the Psychology of Mathematics Education* (pp. 289-294).

TWO METAPHORS FOR REALISTIC MATHEMATICS EDUCATION DESIGN HEURISTICS: IMPLICATIONS FOR DOCUMENTING STUDENT LEARNING

Estrella Johnson
Virginia Tech

The primary goal of this work is to articulate a theoretical foundation based on Realistic Mathematics Education (RME) that can support the analysis of student learning. To do so, I will first frame the guided reinvention and emergent models design heuristics separately in terms of both increasingly general student activity and in terms of concept development. Then, I will consider how the RME design heuristics could inform how one conceptualizes student learning. To do so, I will draw on two metaphors for learning and, by drawing on these two perspectives, propose ways in which the RME design heuristics can inform the analysis of student learning.

Key Words: Realistic Mathematics Education, Learning, Emergent Models, Guided Reinvention, Analytic Methods

Realistic Mathematics Education (RME) is an instructional design theory used to inform the development of inquiry-oriented curriculum. The emergence of such instructional approaches creates a need to investigate student learning in these contexts. However, as it was designed to be an instructional design theory, the current formulations of RME are not articulated in a way that readily supports investigations of student learning. Part of the difficulty in using the current formulations of RME to investigate student learning is due to variations in the ways that the RME design heuristics are discussed in the research literature. In particular, the RME design heuristics are routinely discussed both in terms of increasingly general student activity and in terms of concept development. For instance, both the guided reinvention and emergent models design heuristics are intended to support the creation of a new mathematical reality. However, the creation of a new mathematical reality is discussed both in terms of an activity (Rasmussen, Zandieh, King, & Teppo, 2005) and in terms of object reification (Rasmussen, & Blumenfeld, 2007). As a result, efforts to document the creation of a new mathematical reality (which can be viewed as student learning) are not supported by a clear theoretical foundation.

Building on the work of Johnson (2013), this paper will discuss the RME design heuristics of emergent models and guided reinvention in order to articulate RME in a way that supports analytic techniques for documenting student learning. I will first frame the design heuristics separately in terms of both increasingly general student activity and in terms of concept development. Then, I will consider how the RME design heuristics could inform how one conceptualizes student learning. To do so, I will draw on two metaphors for learning and, by drawing on these two perspectives, propose ways in which the RME design heuristics can inform the analysis of student learning.

RME Design Heuristics in Terms of Activity and Concept Development

Within RME there are a number of heuristics that are meant to guide the design of instruction that supports students in developing formal mathematics by engaging them in mathematical activity. With the *guided reinvention* design heuristic the goal is for "learners to come to regard the knowledge they acquire as their own, personal knowledge, knowledge for which they themselves are responsible" (Gravemeijer & Terwel, 2000, p. 786). This reinvention process can be described by progressive mathematizing, in which students cycle between mathematizing reality (horizontal mathematizing) and mathematizing their own mathematical activity (vertical mathematizing) (Gravemeijer & Doorman, 1999). With the emergent model design heuristic, the goal is to support students' reinvention of mathematics by designing starting point tasks that can elicit informal student strategies that anticipate more formal mathematics. With this heuristic, informal and intuitive *models of* students' mathematical activity transition to *models for* more formal activity (Gravemeijer, 1999).

Captured within both the guided reinvention and the emergent model design heuristics is the duality of engaging in more generalized activity and developing mathematical concepts. By teasing apart these two aspects, two lenses for describing the purpose of these RME design heuristics come into focus. One lens, which considers the two design heuristics in terms of more generalized student activity, places the emphasis on instruction that promotes "socially and culturally situated mathematical practices" (Rasmussen et al., 2005, p. 55). The other lens, which considers the two design heuristics in terms of concept development, places the emphasis on instruction the supports the reification of student activity.

Increasingly General Activity

An emphasis on the students' activity within a given problem context is at the forefront of Rasmussen et al. (2005) discussion of progressive mathematizing, which can be understood as the mechanism supporting guided reinvention. Instead of framing progressive mathematizing in terms of the concepts being developed (e.g., fractions or long division), Rasmussen et al. frame progressive mathematizing in terms of the practices students engage in that promote the evolution of such concepts. As they explain, "this is a nontrivial modification because it calls for attention to the types of activities in which learners engage for the purpose of building new mathematical ideas and methods for solving problems" (p. 55). This focus on student activity is also captured by Zandieh and Rasmussen's (2010) conceptualization of the emergent models heuristic and their definition of a model, where they define a model as "student-generated ways of organizing their activity with observable and mental tools" (p. 58). However, with either design heuristic the point is not merely to design instructional sequences that engage students in mathematical activity. The point is to design instructional sequences that engage students in mathematical activity that is more and more general.

With the guided reinvention heuristic, instruction can be designed with purpose of supporting student activity through progressive mathematizing. During the process of progressive mathematizing, the students' activity shifts repeatedly from horizontal to vertical mathematizing. Initially, horizontal mathematizing is limited to the specific problem context. As students transition to vertical mathematizing, this specific problem context is no longer the focus of the activity, rather the students mathematize their own mathematical activity to support their reasoning in a different or more general situation (Gravemeijer & Doorman, 1999; Rasmussen et al., 2005). Similarly, within the emergent models heuristic, there is an intention to progress students from activity situated within a specific task context to referential, general, and formal activity. In particular, the *model-of/model-for* transition is linked to a shift in the students' activity from referential (where their activity references aspects of the original task setting) to general (where the students activity is no longer tied to the original task setting) (Gravemeijer, 1999). As the students move into general activity, they begin to mathematize aspects of their emerging model. In this way the transition between referential and general activity can be interpreted as the result of vertical mathematizing. Therefore both the guided reinvention and the emergent models heuristics can be framed in terms of increases in the generality of student activity, either as they progress thorough more general layers of activity (in the emergent models heuristic) or as they engage in progressive mathematizing (in the guided reinvention heuristic).

Concept Development

Instead of focusing on the activity in which the students are engaged (and the context in which the student activity is taking place), we could instead focus on the evolution of a mathematical concept. Both the guided reinvention and emergent models heuristics have been connected to reification (Gravemeijer, 1999; Gravemeijer & Doorman, 1999). One can conceive of the guided reinvention and emergent models heuristics as processes through which student activity becomes reified into mathematical objects (Gravemeijer, 1999). This emphasis on reification offers a lens to describe these two design heuristics in terms of the development of the concept, where aspects of the students' mathematical activity become reified as they engage in more general activity.

The guided reinvention heuristic describes this evolution as an expansion of what is experientially real for the students (Gravemeijer, 1999). By engaging in horizontal mathematizing, the students translate aspects of their mathematical reality into mathematical terms. The artifacts of horizontal mathematizing may include inscriptions, symbols, and procedures that represent aspects of an already familiar problem context (Rasmussen et al., 2005). During vertical mathematizing, it is the students' own horizontal mathematizing (and resulting representations/artifacts) that are mathematized. Instead of resulting in representations of an already familiar context (as with the artifacts of horizontal mathematizing), vertical mathematizing results in objects that are now accessible to students on an intuitive level (i.e., these objects are now incorporated into the students' experiential reality). Similarly, the shift from model-of to model-for is related to the process of reification (Gravemeijer, 1999). As students shift from referential activity to general activity "the model becomes an entity in its own right and serves more as a means of mathematical reasoning than as a way to symbolize mathematical activity grounded in particular settings" (p. 164). Therefore, the model – which Gravemeijer (1999) describes as "an overarching concept" (p. 170) – transitions from an artifact of the students' mathematical activity to a mathematical object independent of the students' original activity. Therefore, both the guided reinvention and the emergent models heuristics can be described as processes through which the engagement in progressively more general activity supports the development of mathematical concepts through the reification of student activity.

Framing RME Design Heuristics as Lenses on Student Learning: Two Metaphors

While RME is primarily an instructional design theory, the design heuristics carry with them a view of what it means to learn mathematics. As described by Cobb (2000), RME contends that 1) mathematics is a creative human activity, 2) learning occurs as students develop effective ways to solve problems, and 3) mathematical development involves the creation of a new mathematical reality (p. 317). This suggests that RME can be used to describe the process through which learning takes place. Specifically, learning happens as students engage in activity that is situated in accessible contexts, where this activity brings forth a new mathematical reality.

In the following sections I will discuss how Sfard's (1998) two metaphors for learning, the participation metaphor and the acquisition metaphor, can be used to provide insight into how the design heuristics support student learning. By considering the implication of these two metaphors for learning, I will present two conceptualizations of the notion of a "new mathematical reality". Finally, I will discuss the implications of these two perspectives for analyzing student learning in cases where the instructional design is consistent with the RME design heuristics of guided reinvention and/or emergent models.

Participation Metaphor

Sfard (1998) describes the *participation metaphor* for learning as a view in which "learning" is synonymous with becoming a participant in a community, and "knowledge" is synonymous with aspects of practice/discourse/activity (p. 7). With this view, the emphasis is placed on what the student is *doing*, and the context in which that practice is taking place (as opposed to emphasizing the mental constructs the students have).

Both the guided reinvention and the emergent models heuristics can be framed in terms of increases in the generality of student activity. With this framing, the process of progressive mathematizing (guided reinvention) and the progression thorough more general layers of activity (emergent models) are consistent with Rasmussen et al.'s (2005) notion of *advancing mathematical activity* – where advancing mathematical activity is understood as "acts of participation in different mathematical practices" (p. 53). Therefore, one way to conceptualize student learning in a way that is consistent with an RME perspective is to view student learning as participating in situated activity. Continuing with the participation metaphor, one could ask what it means for student activity (i.e., learning) to support the development of a new mathematical reality. I propose that, from a participation perspective on learning, the creation of a new mathematical reality can be understood as the creation of a new context for further activity, and new ways for students to participate in that context.

When analyzing student learning from a participation perspective, the RME design heuristics provide powerful lenses for documenting student practice and changes in these practices. The various mathematizing activities described in the literature provide examples and characterizations of mathematical practices. Such practices include the mathematizing activities of translating, describing, organizing symbolizing, algorithmatizing, defining, and generalizing (Gravemeijer & Doorman, 1999; Rasmussen et al., 2005; Zandieh & Rasmussen, 2010). Documenting student participation in such practices is a necessary component to documenting student learning. However, it is also necessary to understand changes in the students' practice. The RME design heuristics provide two avenues for analyzing changes in practice. Learning trajectories based on supporting students in progressive mathematizing and/or progressing through layers of generality provide a framework for analyzing how the mathematical practices of the students are changing in regards to the generality of their activity. Additionally, the notion of a new mathematical reality provides a way to discuss both changes in the context of the students' activity and changes in how students participate in this new context.

Acquisition Metaphor

With the *acquisition metaphor*, learning is viewed as the acquisition of knowledge and concepts. This perspective "makes us think about the human mind as a container to be filled with certain materials and about the learner as becoming an owner of these materials" (Sfard, 1998, p. 5). Therefore, this perspective places the emphasis on concept development, where "concepts are to be understood as basic units of knowledge that can be accumulated, gradually refined, and combined to form ever richer cognitive structures" (p. 5).

This perspective comes to the forefront when the guided reinvention and emergent models heuristics are framed in terms of reification. With guided reinvention, mathematical concepts develop as a result of horizontal and vertical mathematizing. During vertical mathematizing, the students mathematize their own horizontal mathematizing (and resulting representations/artifacts). This results in the students' activity becoming a new type of object that is accessible to them on an intuitive level. Similarly, the emergent models heuristic describes a process through which students' activity emerges first as a *model-of* the students' informal activity and then transitions to being a *model-for* supporting students' more formal reasoning (and thus an object in the students' mathematical reality). From this perspective, learning not only supports the creation of a new mathematical reality (as it did with the participation metaphor), learning can be viewed as synonymous with the creation of a new mathematical reality. I propose that, from an acquisition perspective on learning, the creation of a new mathematical reality can be conceptualized as the incorporation of new mathematical objects into the students' experiential reality. These new mathematical objects can be understood as concepts that form "ever richer cognitive structures" (Sfard, 1998, p. 5), and the fact that they become incorporated into the students' experiential reality reflects that the students are able to access these concepts on an intuitive level.

In order to document student learning from an acquisition perspective the focus must be on the development of the mathematical concepts. With both the emergent models and guided reinvention design heuristics, the mathematical concepts develop as aspects of the students' mathematical activity become reified. Instead of considering the reification of a global concept, we can consider a smaller grain size of analysis by discussing the documentation of local evidence of student learning. For the emergent model construct, local changes can either be 1) related to the form of the model the, as described by the chains of signification construct (Gravemeijer, 1999), or 2) related to the function of the model, as described by the record-of/tool-for construct (Larsen, 2004). Documenting such local shifts may include looking for indications that one sign has slid under a subsequent sign and looking for indications that a record-of student activity is serving as a tool-for subsequent student activity. Both of these local shifts support the reification of the global model (i.e., student learning from an acquisition perspective). In the case of the guided reinvention heuristic, the goal is to find evidence of an expansion in what is experientially real for the students. These additions reflect that aspects of the students' activity have become objects that are now accessible for further mathematizing. From an acquisition perspective, this is understood as a creation of a new mathematical reality, where new mathematical objects become incorporated into the students' experiential reality.

Conclusion

This paper was written to explore the implications of RME for documenting student learning. Both the guided reinvention and emergent models design heuristics support the development of new mathematical realities by engaging students in increasingly generalized activity, and both can be described either in terms of more generalized activity or in terms of concept development. By focusing independently on these two aspects of the design heuristics, I was able to draw on Sfard's (1998) participation and acquisition metaphors for learning in order to discuss how these design heuristics support student learning.

Considering the design heuristics in light of these two perspectives on learning afforded a powerful lens for making sense of the idea of a new mathematical reality and for discussing what could be considered as evidence for student learning. I propose that, from a participation perspective, the creation of a new mathematical reality can be understood as the creation of a new context for further activity, and new ways for students to participate in that context. From a participation perspective, the RME design heuristics suggest a number of ways to document student learning. This includes: documenting the mathematizing activities that students are engaged in, documenting how the mathematical practices of the students are changing in terms of the generality of their activity, and documenting changes in the students mathematical reality – both in terms of the context of the students' activity and in terms of how students participate in this new context. From an acquisition perspective on learning, I propose that the creation of a new mathematical reality can be conceptualized as the incorporation of new mathematical objects into the students' experiential reality. The incorporation of these new objects reflects that they have become accessible to students on an intuitive level. Again, the RME design heuristics suggest a number of ways to document student learning from an acquisition perspective. This includes: documenting when that one sign has slid under a subsequent sign, documenting when a record-of student activity is serving as a tool-for subsequent student activity, and documenting incremental additions to the students' mathematical reality.

References

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (Vol. Lawrence Erlbaum Associates, pp. 307-334). Mahwah, NJ.

Gravemeijer, K. (1999). Emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1*(2), 155-177.

Gravemeijer, K., & Doorman, M. (1999). Context problems in Realistic Mathematics Education: A calculus course as an example. *Educational Studies in Mathematics, 39*, 111-129.

Gravemeijer, K., & Terwel, J. (2000). Hans Freudenthal: A mathematician on didactics and curriculum theory. *Journal of Curriculum Studies 32*(6), 777-796.

Johnson, E. (2013). *Implications of Realistic Mathematics Education for Analyzing Student Learning* . Sixteenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education. Denver, CO.

Rasmussen, C., & Blumenfeld, H. (2007). Reinventing solutions to systems of linear differential equations: A case of emergent models involving analytic expressions. *Journal of Mathematical Behavior, 26*, 195-210.

Rasmussen, C., Zandieh, M., King, K., & Teppo, A. (2005). Advancing mathematical activity: A practice-oriented view of advanced mathematical thinking. *Mathematical Thinking and Learning 7*(1), 51-73.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4-13.

Zandieh, M., & Rasmussen, C. (2010). Defining as a mathematical activity: A framework for characterizing progress from informal to more formal ways of reasoning. *Journal of Mathematical Behavior, 29*, 57-75.

HOW TO MAKE TIME: THE RELATIONSHIPS BETWEEN CONCERNS ABOUT COVERAGE, MATERIAL COVERED, INSTRUCTIONAL PRACTICES, AND STUDENT SUCCESS IN COLLEGE CALCULUS

Estrella Johnson
Virginia Tech

Jessica Ellis and Chris Rasmussen
San Diego State University

This report draws on data collected by the Characteristics of Successful Programs in College Calculus project in order to investigate issues around coverage and pacing. This includes identifying what topics are being taught in Calculus I, determining the extent to which instructors and departments feel pressure to cover a set amount of material, and investigating possible relationships between concerns over coverage, instructional practices, and the nature of the material covered at five institutions selected for having successful Calculus programs.

Key Words: Calculus, Coverage, Teaching Practices, Pacing

The Characteristics of Successful Programs in College Calculus (CSPCC) project is a large empirical study designed to investigate Calculus I across the United States. While Calculus I is offered at nearly every college and university across the nation, and taken by approximately 300,000 students every fall, prior to CSPCC very little data had been collected about what happens in Calculus I (Bressoud et al., 2013). The primary focus of the CSPCC project is to identify factors that contribute to student success and understand how these factors are leveraged within highly successful programs. However, in addition to addressing these primary research goals, the CSPCC project has also amassed a wealth of data on the nature of Calculus I courses across the nation. In this report, we aim to draw on the CSPCC data in order to investigate issues around coverage and pacing. This includes investigating what topics are being taught in Calculus I and determining the extent to which instructors and departments feel pressure to cover a set amount of material. Further, because concerns over coverage are often cited as reasons to not implement reform-oriented instructional practices (Christou et al., 2004; Johnson et al., 2013; McDuffie & Graeber, 2003; Wagner, Speer, & Rossa, 2007; Wu, 1999) we will investigate relationships between teaching methods and concerns over coverage.

Theoretical Background

Students are citing poor instruction in their mathematics and science courses, with calculus instruction and curriculum often singled out, as a contributing reason for why they are discontinuing in STEM fields (Rasmussen & Ellis, 2013; Seymour, 2006; Thompson et al., 2007). Some specific problems with their learning experiences that students identified include: courses that were over-stuffed with material; pacing that inhibited comprehension and reflection; not including applications or conceptual discussions; and "faculty modes of teaching that suggested that they took little responsibility for student learning" (Seymour, 2006, p. 4). Thus, as reported by students, shallow treatments of large amounts of material and unresponsive teaching strategies are contributing to their reasons for leaving STEM majors.

The response from teachers seems to be that pressure to cover a set amount of material precludes efforts to adopt reform-oriented teaching strategies. For instance, in a case study of two mathematicians trying to implement reform curriculum in mathematics courses for preservice teachers, McDuffie and Graeber (2003) identified a number of institutional norms and policies that curtailed the mathematicians' efforts. As stated by one of the mathematicians:

> If you've got courses that link together, as most of the math curriculum does…there's an expectation that a certain amount of material be covered… And so you're fighting this constant battle…It means that you're limited on how much time you can spend to do real constructivist activities where the depth of knowledge is really greater (McDuffie & Graeber, 2003, p. 336).

Wu (1999) echoed this sentiment. In an op-ed reaction to mathematics education reform, he proposed that, "if the amount of material to be covered in a course can be greatly reduced … and students are expected to spend 8 years in college… then we can all safely abandon the lecture format and engage in a wholesale application of the guide-on-the-side philosophy" (p. 4). As examples of deliberate reduction in the material to be covered, Wu offers the textbooks *Calculus* by Hughes-Hallet et al. and *Calculus Concepts* by La Torre et al.

Taken as a whole, these reports from teachers and students suggest that 1) calculus courses are overburdened with content, and 2) in order to cover such large amounts of material teachers cannot implement reform-oriented instruction. In this study, we draw on the CSPCC data to investigate the validity of these claims using data collected at 197 research universities across the nation, including five institutions that have been selected for having particularly successful[1] Calculus I programs. Specifically, we investigate the following question: *In the PhD granting institutions with successful calculus programs, what is the relationship between concerns about coverage, instructional practices, and the nature of the material covered?*

Embedded in this question are issues regarding the expectations of students and faculty. These expectations relate to who is responsible for learning, where learning occurs, and how much material is reasonable to cover. Theoretically, we see these types of expectations as part of the didactical contract (Brousseau, 1997). The notion of didactical contract refers to the set of reciprocal expectations and obligations between the instructor and the students, most of which are implicitly formed through patterns of interaction. For example, at the secondary school level students do not expect to have to cover large amounts of material on their own at home. Much of learning therefore occurs in class and students and their teacher are mutually responsible for learning. At the university level, however, these expectations and obligations may shift – the amount of material covered increases, instructors tend to lecture more compared to secondary school teachers, and instructors expect students to learn more on their own at home. Students are often left feeling that their calculus course is overstuffed and taught in an uninspiring and unresponsive manner (Seymour, 2006). It is precisely these aspects of the didactical contract that we aim to unpack at institutions with more successful calculus programs.

Research methodology

In order to answer our research question, we draw on data collected in the two phases of the CSPCC project. The first phase of the CSPCC study involved surveys sent to a stratified random sample of students and their instructors at the beginning and the end of Calculus I. These surveys were designed to gain an overview of the various calculus programs nationwide, and to determine which institutions had more successful calculus programs. Here success was defined by a combination of student variables: persistence in Calculus as marked by stated intention to

[1] Our measures of success are outlined in the "Research methodology" section.

take Calculus II; affective changes, including enjoyment of math, confidence in mathematical ability, and interest to continue studying math; and passing rates. In the second phase of this project, surveys were analyzed in order to select four or five successful schools of each type (community college, Bachelor's granting, Master's granting, and Doctoral granting). We then conducted three-day site visits at each of the 18 institutions selected, during which we: interviewed students, instructors, and administrators; observed classes; and collected exams, course materials, and homework.

To understand departmental[2] *concerns about coverage*, we drew on instructors' agreement levels to the following survey prompts: When teaching my Calculus class, I (a) had enough time during class to help students understand difficult ideas, and (b) felt pressured to go through material quickly to cover all the required topics. To understand the departmental *instructional practices*, we drew on instructors' reports or the frequency of 8 instructional activities: (a) show students how to work specific problems; (b) have students work with one another; (c) hold a whole-class discussion; (d) have students give presentations; (e) have students work individually on problems or tasks; (f) lecture; (g) ask questions; and (h) ask students to explain their thinking. For both sets of questions, instructors were prompted to provide a response ranging from 1 to 6 on a Likert scale, with 1 meaning "not at all" and 6 meaning "very often". Descriptive and correlational analyses were conducted on these questions, with results discussed below.

There were 238 instructors who answered the above questions, 50 of who came from one of the five selected Doctoral granting institutions: Western Religious University (WRU), Northern Tech (NT), University of West Coast State (UWCS), University of Northern State (UNS), New England Polytechnic Institute (NEPI). Table 1 provides a brief overview of these institutions.

Table 1. Summary of selected institutions

Doctoral Institution	**Instructors with survey responses**	**Term length (weeks)**	**Text Used**
Western Religious University (WRU)	3	15	*Single Variable Calculus: Early Transcendentals* by Stewart
Northern Tech (NT)	7	14	*Calculus, Single and Multivariable (Fifth Edition)* by Hughes-Hallett, et al.
University of West Coast State (UWCS)	4	11	*Calculus: Early Transcendentals* by Jon Rogawski
University of Northern State (UNS)	30	15	*Calculus, Single and Multivariable (Fifth Edition)* by Hughes-Hallett, et al.
New England Polytechnic Institute (NEPI)	6	7	*Calculus: Early Transcendentals (7th edition)* by Edwards and Penny

[2] For the purposes of this analysis, we consider instructor responses together as representative of department concerns about coverage and instructional practices. In later analyses we consider in depth the variation among instructors within departments.

To understand the *nature of the material covered*, course syllabi and the departmental course list of required sections to be covered were analyzed. A master list of section titles was sorted into five categories: function review, limits, derivatives, differentiation rules, applications of differentiation, and integrals. Equivalent section titles were then grouped together to better reflect commonalities between the topics. For instance, the sections entitled *The Derivative as a Rate of Change, Rates of Change*, and *The Derivate and Rates of Change* were condensed into one heading.

Results

To answer our research question, we first report on departmental concerns about coverage; departmental teaching practices; and, for the five selected institutions, the material intended to be covered. We then present on how each these are related to one another.

Departmental concerns about coverage

As shown in Table 2, there are no significant differences between how concerned the departments are about coverage. On average, instructors at both the selected and not selected institutions reported having enough time to help students understand difficult ideas with around 4/6 frequency, and reported feeling pressured to go through the material quickly to cover all the required topics around 3/6 frequency.

Table 2. Departmental reports of concern for coverage at selected and not selected institutions.

When teaching my Calculus class, I: *(1=Not at all; 6=Very often)*	**Not Selected**	**Selected**
had enough time during class to help students understand difficult ideas.	4.19 (1.31)	4.42 (1.14)
felt pressured to go through material quickly to cover all the required topics.	3.06 (1.46)	3.33 (1.35)

Note.* = $p \leq .10$, ** = $p \leq .05$, *** = $p \leq .001$; Std. dev. is in parentheses.

Departmental instructional practices

As shown in Table 3, there are significant differences between the reported instructional practices of the instructors at the selected and not selected institutions. Specifically, instructors at the five selected institutions report higher frequencies of having students work with one another, holding a whole-class discussion, having students give presentations, and asking students to explain their thinking.

Table 3. Instructor reports of instructional activity at selected and not selected institutions.

During class time, how frequently did you: *(1=Not at all; 6=Very often)*	**Not Selected**	**Selected**
show students how to work specific problems?	5.14 (1.12)	5.22 (.89)
have students work with one another? ***	2.71 (1.65)	4.28 (1.84)
hold a whole-class discussion? **	2.68 (1.56)	3.32 (1.66)
have students give presentations? ***	1.47 (.91)	2.35 (1.74)

have students work individually on problems or tasks?	2.82 (1.60)	3.18 (1.66)
lecture?	5.25 (1.20)	5.12 (1.17)
ask questions?	5.15 (1.08)	5.08 (1.09)
ask students to explain their thinking?**	3.77 (1.50)	4.30 (1.42)

Note.* = p ≤ .10, ** = p ≤ .05, *** = p ≤ .001; Std. dev. in parentheses.

Nature of material covered at selected institutions

Analysis of the common syllabi from the five selected institutions identified six areas that were included in at least one of the Calculus I programs: Function Review, Limits, Derivatives, Differentiation Rules, Differentiation Applications, and Integrals. Only two of the schools, UWCS and UNS, covered sections in all six areas. Table 4 shows the number of sections in each area that the five schools included in their Calculus I course as well as their pace (number of topics per week). Notice that WRU did not include any review sections, NT did not include any sections on limits, and NEP did not cover any sections on integration.

Table 4. Nature of material covered at selected institutions

Topic covered	**Institution**				
	WRU	**NT**	**UWCS**	**UNS**	**NEPI**
Function Review	0	6	3	6	5
Limits	4	0	8	2	4
Derivatives	2	4	3	6	1
Differentiation rules	4	7	7	7	5
Differentiation Applications	7	4	8	7	13
Integrals	6	8	6	5	0
Total	***23***	***29***	***35***	***33***	***28***
Pacing (Topics per weeks in term)	**1.53**	**2.07**	**3.18**	**2.20**	**4.00**

Differences were also found within the main areas. In total, syllabi from the five schools included 84 different sections. However, only 7 topics were common to at least four of the five schools. These topics were: Limits and Continuity, Differentiation Rules (power, sum, product, quotient, exponential, chain, trigonometric), Related Rates, Max/Min/Optimization, Optimization and Modeling, Linear Approximations, and The Fundamental Theorem of Calculus. Additionally, there was variability among the sections that defined the derivative, both in terms of the number of sections covered and in terms of the topics. In this area sections names included: The derivative as the slope of a tangent line, The derivative as a rate of change, Derivative at a point, Derivative Function, and Definition of the Derivative. Finally, the schools varied greatly in the pace at which they went through sections, ranging from 1.53 sections per week at WRU to 4 sections per week at NEPI.

Relationship between concerns about coverage and nature of the material covered at the selected institutions.

In order to understand the relationship between departmental concerns and the nature of the material covered at the selected institutions, we first conducted correlation analysis between instructors' responses to the two questions regarding concerns about coverage and the intended pacing as determined in the syllabi analysis. There is no correlation between (a) departmental reports of having enough time during class to help students understand difficult ideas and intended pacing, $r(48) = .053$, $p = .713$; or between (b) feeling pressured to go through material quickly to cover all the required topics and intended pacing, $r(48) = .070$, $p = .632$.

We then looked in depth at the two schools with the largest difference in the number of sections covered per week: WRU with 23 sections included in their required section list to be covered in a 15-week term (1.53 topics per week) and NEPI with 28 sections in their departmental syllabi to be covered in a 7-week term (4 topics per week). When asked if they felt that they had enough class time to help their students understand difficult ideas, 2 of the 3 teachers from WRU responded that they *did not* feel like the had enough time (both answering with a 2 out of 6 on a Likert scale with 1 being not at all and 6 being very often). When answering the same question, only 1 of the 6 NEPI instructors gave a rating of 3 or less. Additionally, when asked if they felt pressured to go through material quickly to cover all the required topics, all 3 of the WRU replied with a score of 4 or more (again on a Likert Scale with 1 being not at all and 6 being very often). For the same question, 4 of the 6 NEPI replied with a score of 4 or more. These findings indicate that while instructors at NEPI (the institution with the quickest pace) felt pressured to go through the material quickly, they also felt like they had time to help their students understand difficult ideas. Conversely, instructors at WRU (with the slowest pace) felt both pressured to quickly cover the material and like they did not have enough time to help their students understand difficult topics.

Relationship between instructional practices and the nature of the material covered at the selected institutions.

To understand the relationship between departmental instructional practices and the nature of the material covered, we again first conducted correlational analyses between the eight reported instructional practices and the intended pacing at each of the five selected institution. Of the eight instructional practices, only one is correlated to pacing. There is a strong negative correlation between the frequency that students worked together and the intended pacing, $r(48) = -.548$, $p < .001$. This result implies that instructors who cover material quickly do not have students work in groups often. Indeed, instructors at the two institutions with the quickest pacing, NEPI (4 sections per week) and UWCS (3.18), reported that they rarely had students work in groups. However, at both UNS (2.2 section per week) and NT (2.07 sections per week) the majority of instructors reported that they often had students work in groups.

Relation between concerns about coverage and instructional practices at selected and not selected institutions.

Lastly, we looked at the relationship between reported departmental concerns about pacing and instructional practices. Again, we conducted correlational analyses between the two questions regarding concerns about coverage and the eight questions regarding instructional practices. Among the five selected institutions, there was a slight positive correlation between having enough time during class to help students understand difficult ideas and the frequency of the instructor asking the students questions, r(48) = .268, $p = .060$, and between feeling pressured to go through material quickly to cover all the required topics and the frequency of lecture, r(47) = .263, $p = .068$. These results suggest that, at the selected institutions, instructors who reported having enough time to help their students with difficult ideas often asked their students many questions during class, and those instructors who felt pressured to rush through material quickly tended to lecture more.

Among the not selected institutions, there was a strong correlation between having enough time during class to help students understand difficult ideas and the frequency of showing students how to work specific problems, r(186) = .296, $p < .01$, and a slight correlation between having enough time during class to help students understand difficult ideas and the frequency of having students give presentations, r(184) = .134, $p = .068$. Additionally, there were strong negative correlations between feeling pressured to go through material quickly to cover all the required topics and showing students how to work specific problems, r(182) = -.204, $p = .006$, and having students give presentations, r(180) = -.182, $p = .014$. These results indicate that, at the institutions not selected, instructors who reported having enough time to help their students with difficult ideas often showed their students how to work specific problems and had them give presentations. Further, instructors who felt pressured to rush through material quickly tended to infrequently do these activities.

Discussion

Given that these five institutions were selected based on student success (including persistence in Calculus, positive affective changes, and high pass rates), these results may suggest components of didactical contracts that support student success. For instance, between the selected and not selected institutions, there were no differences in the amount of time instructors felt like they had to help students through challenging material. However, there was a difference with how the instructors chose to use their time. When instructors report having enough time to help student understand difficult material, instructors at the selected institutions are more likely to use that time asking their students questions during class and instructors at the not selected institutions are more likely to use that time showing their students how to work specific problems and having them give presentations. Additionally, between the selected and not selected institutions, there were no differences in the amount of pressure that instructors felt to cover material (and in fact, when looking at the five selected institutions, there is no correlation between the reported concerns about coverage and the intended pacing of the course). There was, however, a difference in how instructors at the selected and not selected institutions chose to cover the material. Instructors at the selected institutions reported higher frequencies of having students work with one another, holding a whole-class discussion, having students give presentations, and asking students to explain their thinking. Thus, at these selected institutions it appears that part of the didactical contract between the instructors and their students involves covering material, sometimes large amounts, in ways that will involve and engage students in their learning.

References

Bressoud, D. (2009). Is the Sky Still Falling? Notices of the AMS, 56:2–7.

Bressoud, D., Carlson, M., Mesa, V., Rasmussen, C. (2013). The Calculus Student: Insights from the MAA National Study. International Journal of Mathematical Education in Science and Technology. 44. 685–698.

Brousseau, G. (1997). *Theory of didactic situations in mathematics* (N. Balacheff, Trans.). Dordrecht, Netherlands: Kluwer.

Christou, C., Eliophotou-Menon, M., & Philippou, G. (2004). Teachers' concerns regarding the adoption of a new mathematics curriculum: An application of CBAM. *Educational Studies in Mathematics, 57*(2), 157-176.

Hurtado, S., Eagan, M. K., & Chang, M. (2010). Degrees of success: Bachelor's degree completion rates among initial STEM majors. Retrieved from http://www.heri.ucla.edu/nih/downloads/2010%20%20Hurtado,%20Eagan,%20Chang%20-%20Degrees%20of%20Success.pdf

Johnson, E., Caughman, J., Fredericks, J., & Gibson, L. (2013). Implementing inquiry-oriented curriculum: From the mathematicians' perspective. *Journal of Mathematical Behavior. 32 (4).* 743 - 760

McDuffie, A.R., & Graeber, A.O. (2003). Institutional norms and policies that influence college mathematics professors in the process of changing to reform-based practices. *School Science and Mathematics*, 103(7), 331-344.

Rasmussen, C., & Ellis, J. (2013). Who is switching out of calculus and why? In Lindmeier, A. M. & Heinze, A. (Eds.). *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 4 (pp. 73-80). Kiel, Germany: PME.

Seymour, E. (2006), Testimony offered by Elaine Seymour, Ph.D., University of Colorado at Boulder, to the Research Subcommittee of the Committee on Science of the U.S. House of Representatives Hearing on Undergraduate Science, Math and Engineering Education: What's Working? Wednesday, March 15, 2006

Thompson, P. W., Castillo-Chavez, C., Culbertson, R. J., Flores, A., Greely, R., Haag, S., et al. (2007). Failing the future: Problems of persistence and retention in science, technology, engineering, and mathematics majors at Arizona State University. Tempe, AZ: Office of the Provost.

Wagner, J. F., Speer, N. M., & Rossa, B. (2007). Beyond mathematical content knowledge: A mathematician's knowledge needed for teaching an inquiry oriented differential equations course. *Journal of Mathematical Behavior*, 26, 247–266.

Wu, H. (1999). The joy of lecturing - With a critique of the romantic tradition of education writing. In S. G. Krantz (Ed.) *How to teach mathematics* (pp. 261-271). Providence, RI: American Mathematical Society.

75

THREE CONCEPTUALIZATIONS OF THE DEFINITE INTEGRAL IN MATHEMATICS AND PHYSICS CONTEXTS

Steven R. Jones, Brigham Young University, Provo, UT

Student understanding of the integral is a topic of recent interest in undergraduate education. We are just beginning to learn how different interpretations of the definite integral influence student thinking in both mathematics and science classrooms. This paper examines the relative "productivity" of three conceptualizations of the definite integral in mathematics and physics tasks. It appeared that the notion of the integral as an "addition over many pieces" was especially useful for understanding applied problems.

Key Words: Integration, Calculus, Summation, Accumulation, Physics and engineering

Student understanding of the integral is a valuable topic, since students who continue into further calculus courses will encounter the integral often (Stewart, 2012; Thomas, Weir, & Hass, 2009) and since integration serves as the basis for many real world applications (Beichner, 1994; Christensen & Thompson, 2010; Pollock, Thompson, & Mountcastle, 2007). The integral is used to define and compute various quantities in physics and engineering (Hibbeler, 2012; Serway & Jewett, 2008). However, an overreliance on certain interpretations of the integral, such as an "area under a curve," might limit the integral's applicability to these other areas (Jones, 2013; Sealey, 2006). Improved application of the integral to science classes should be of primary concern for calculus instructors due to its nature as a service course and the large portion of science students enrolled in them (Bressoud, 2012; Mustoe, 2002).

Calculus textbooks primarily speak of the definite integral in three ways: (a) the area under a curve, (b) the difference in values of an anti-derivative, and (c) the limit of Riemann sums (see Stewart, 2012; Thomas et al., 2009). Recent research has also shown some additional meanings and nuances students give to these three notions. Jones (2013) discusses (a) students incorporating the ideas of a perimeter, created by the integrand, differential, and limits of integration; (b) a function-matching game where the integrand is thought to have come from some other "original function;" and (c) an addition over infinitely many infinitesimally small pieces. Hall (2010) also shows how everyday language can influence how students interpret words such as "definite" integrals. Sealey and others investigated how students understand Riemann sums and incorporate them into their understanding of limits (Engelke & Sealey, 2009; Sealey & Oehrtman, 2005).

Yet none of these studies attempts to deeply analyze how these various conceptualizations of the integral play out in understanding expressions and equations involving the definite integral in mathematics and physics settings. This paper contributes an analysis of the "productivity" of these three interpretations of the definite integral in both mathematics and physics contexts.

Symbolic Forms of the Integral and Framing

Symbolic forms (Sherin, 2001) are a subset of *cognitive resources* (Hammer, 2000), which are elements of cognition that are accessible by an individual as a unit. However, unlike the traditional idea of a unitary "concept" within a person's cognition, a cognitive resource may be made up of other, smaller resources (Hammer & Elby, 2002). As an example, a student's *concept* of integration might not be a single entity, but may rather be made up of smaller units, including ideas of area, summations, functions, or differentials. Each of these may be further made up of even smaller units, such as lines, perimeters, rectangles, parts of a whole, the power rule in reverse, and so forth. If this is the case, one cannot claim that a student's *concept* of integration is one fixed cognitive object that is either correct or incorrect (see Hammer, 2000). Thus certain interpretations of the integral may turn out to be useful and productive in one context, but turn out *not* to be useful nor productive in another.

A symbolic form (Sherin, 2001) is a specific type of cognitive resource consisting of a blend (Fauconnier & Turner, 2002) between two components: a *symbol template* and a *conceptual schema*. A symbol template is merely the arrangement of the symbols in an expression or equation. In this paper, the symbol templates under consideration are that of the *definite integral*, $\int_{[]}^{[]}[]d[]$ or $\int_{[]}[]d[]$, where each "box" can be filled with a symbol. A conceptual schema, on the other hand, refers to the meaning underlying the symbols. Students have demonstrated several ways to assign meaning to these symbols, including *function matching*, *perimeter and area*, and *adding up pieces* (Jones, 2013), which correlate with the (a) values of an anti-derivative, (b) area under a curve, and (c) limit of Riemann sums conceptualizations, respectively. Note that even though the anti-derivative notion is often thought of as the province of *indefinite* integrals, students showed a strong tendency to conceptualize *definite* integrals through the anti-derivative lens as well (Jones, 2013). A brief description of these three symbolic forms is now provided.

Function matching: This symbolic form may be considered the reification of the anti-derivative process into an object (Sfard, 1991). Students give the integrand the meaning of having come from some other "original function," which *became* the integrand through a derivative (Jones, 2013). The meaning of the integral is a matching game, trying to get back to the original function. The limits of integration refer to the "competing terms" (Sherin, 2001) of values of an anti-derivative.

Perimeter and area: This symbolic form is related to the conception of an integral as an area under a curve. Students imbue each "box" in the symbol template with the meaning of being part of a perimeter of a shape in the plane (Jones, 2013). The limits of integration are represented by vertical lines for the two sides and the differential dictates which axis serves as the "bottom" of the shape. In this conception, the region of interest is *not* divided up, but is considered as a static whole, often represented by shading in the entire area all at once.

Adding up pieces: This third symbolic form bears a resemblance to the Riemann sum. The region of interest is conceptually held to be divided into "tiny pieces," which are often thin rectangles. Generally, only one of these pieces, called a *representative rectangle* (Jones, 2013), is used to analyze the integral and determine its properties (though "rectangle" can be expanded to include any shape depending on the integral). The quantities represented by these small pieces are systematically added up in a dynamic fashion with a "starting" piece and an "ending" piece. Interestingly, most students thought of there being an infinite number of tiny pieces, requiring an infinite summation over those infinite number of pieces (Jones, 2013).

Since students may hold several different notions of the integral in their cognition, a tacit "choice" must be made about which conceptualization to draw on for a given task. This choice is made through *framing* (MacLachlan & Reid, 1994), which means "a set of expectations an individual has about the situation in which she finds herself that affect what she notices and how she thinks to act" (Hammer, Elby, Scherr, & Redish, 2005, p. 97). Thus, a student's expectations regarding the interview context, the specific task they are working on, and what counts as a "good answer" all affect which cognitive resources the student might employ for a given problem. Fortunately, this allows for the opportunity to determine how useful or productive certain conceptualizations of the integral are, depending on the context.

Data Collection and Analysis

Interviews were conducted with eight students selected from a major university on the West Coast of the United States. The students were each interviewed twice, in pairs. The first interview consisted of "mathematics-framed" tasks, making use of integrals similar to those found in typical calculus texts (Stewart, 2012; Thomas et al., 2009). The second interview, which happened one week later, consisted of "physics-framed" tasks, using integrals like those found in calculus-based physics and engineering texts (Hibbeler, 2012; Serway & Jewett, 2008).

The students discussed their responses to the tasks with each other at the board until both participants were satisfied. The videotaped interview sessions and the researcher's notes were the primary sources of data for the study. The data were analyzed by determining locations in the interviews where students demonstrated usage of a particular symbolic form, and then scrutinizing the data, in several iterations, for confirming or disconfirming evidence of that symbolic form (see Strauss & Corbin, 1998). Once a student's usage of a symbolic form appeared to have secure footing, the form was contextually examined for evidence of its effects on student thinking in mathematics-framed and physics-framed settings.

Productivity in Mathematics-framed and Physics-framed Contexts

In this section I discuss the relative "productivity" of the three conceptions described in the previous section in both mathematics and physics contexts. Here, "productivity" means the ability of a symbolic form to facilitate an understanding of the integral that is satisfactory to the student and generally sound in its relationship to commonly accepted notions of the integral.

Function Matching

One student, who I call Darius, consistently favored the *function matching* symbolic form in his work. By relying on this type of thinking, Darius had a well-conceived motivation for *why* he could calculate pure mathematics integrals in the way he did. After working out the solution to $\int_1^2 \left(2/x^3 - x^2\right) dx$ with his partner, he offered this explanation for what the integral meant.

> *Darius*: In an integration the *dx* is always essential, because it shows that this entire thing [waves hand over the integrand, "$2/x^3 - x^2$"] is a derivative of x… The fact that this entire thing is sitting right next to each other, and *dx* outside, means that basically this entire function [motions hand over "$2/x^3 - x^2$"] is the derivative of an original function.

Darius used the idea of trying to recover an "original function" that became the integrand via a derivative as the key motive for why the rules of differentiation needed to be done "in reverse" to determine the solution. If a derivative yielded an "x^2" then an integral needed to figure out what function would turn into x^2 under differentiation. In this way, the *function matching* symbolic form was *productive* for Darius in the pure mathematics context by allowing him to feel satisfied with his reasoning for the procedures he carried out.

However, by contrast, the *function matching* conceptualization failed to be as effective in the physics-framed context, where multivariate functions are often encountered. Carlos and Curtis attempted to employ this thinking to understand the integral equation $F = \int_S P\, dA$, in which the pressure over a surface, S, is used to calculate the total force exerted on S. The students began their discussion by talking about how they could go about calculating the answer through anti-derivatives.

Carlos: Like, will that [points to P] be a function of dA?

Curtis: That means P has to be a function of, uh, with respect to A, I guess. Because… yeah, so we would have to be given, in order to solve this with actual values, we would need to be given P, pressure, as a function of area… [quietly to Carlos] How would you say that? [Long pause.]

By focusing on creating an anti-derivative that "matches" with the integrand P, these students were led to assume that P had to be a function of A since the differential indicates *how* to take the anti-derivative (Jones, 2013). But this would lead to a meaningless express, such as $P = A^2 + A$ for example, which does not accurately represent the physics context, since P is *not* a function of A. Hence, this conceptualization proved *less productive* for understand physics integrals.

Perimeter and Area

During the mathematics-framed interview, Carlos and Curtis often relied on the *perimeter and area* symbolic form to understand the meaning of the integrals. They solved the integral $\int_1^2 \left(2/x^3 - x^2\right) dx$ by splitting it up into $\int_1^2 2/x^3\, dx$ and $\int_1^2 x^2\, dx$ and subtracting the results. Curtis used the "area under a curve" thinking to explain why this was a valid approach.

Curtis: Like, if we have two curves [draws one curve with a second one below it], instead of having a single integral to solve this total area all at once, we're finding the integral of the top one [spreads hands from the upper curve to the horizontal axis] and then we're subtracting this area [outlines shape from the lower curve to the horizontal axis].

The conceptualization of the integral as an area under a curve allowed Curtis to understand this property of integrals, namely splitting up over subtraction. He could appeal to areas to intuitively explain why the resulting two integrals were equal to the original. By "removing" the portion below the lower curve from the overall area under the higher curve, he was able to justify why this calculates the area *in between* the curves. Thus, this conception was *productive* for him in understanding pure mathematics integrals.

Yet, like *function matching*, the *perimeter and area* notion was less effective in the physics context. For example, Brian struggled to use this reasoning to interpret the meaning of the integral equation $F = \int_S PdA$. Brian drew a one-dimensional curve in the plane, labeled it P, labeled the horizontal axis A, and marked off vertical lines for the left and right sides of a bounded region. However, after Brian had produced this graph, he appeared unsatisfied and did not know how to use his picture to explain the integral.

Brian: So, if we took the integral of that, it would be, it would be all this… [shades in the region underneath the P-curve]. And that would be the total force it would exert… umm… I feel like I skipped a step.

Brian was at a loss as to how to reconcile his use of A as an independent variable on the horizontal axis with the fact that the area of the surface should not be changing. Furthermore, he seemed uncomfortable in describing why the area of the region he had just created should even represent the total force at all, stating that he felt like he "skipped a step." Thus, this conceptualization of the integral was *less productive* for making sense of the physics equation.

Adding Up Pieces

By comparison, students drawing on the *adding up pieces* symbolic form not only understood pure mathematics integrals, but were also able to satisfactorily explain a variety of applied physics integrals. Since the Riemann sum, which consists of an addition of values over many (small) pieces, is used as the basis for a common mathematics definition of the definite integral, $\lim_{\Delta x \to 0} \sum_{k=1}^{n} f(x_k)\Delta x$, it suffices to say that the *adding up pieces* conception is productive in the pure mathematics context. Instead of focusing on the mathematics context, I focus here on how Curtis and Brian switched from *function matching* and *perimeter and area* to *adding up pieces,* which enabled them to push past their difficulties in making sense of the physics equation $F = \int_S PdA$. Curtis began switching from *function matching* to *adding up pieces* by drawing a rectangular surface, *S,* and then drawing a small *representative square* inside of it.

Curtis: So, force equals pressure times area. We have the area [points to the small square] and we have the pressure function [makes "incoming" motion toward the square]… And then we have pressure times area, so we're actually finding… force… So, we're integrating over these infinitesimally small [pieces], which each one composes of a force, so we're integrating force, and adding up all the infinitesimally small pieces of force, to find the total force.

Similarly, after Brian pondered his "skipped step" for a while, I told him to think of the table I was sitting at as *S* and asked him if the integral applied to that situation. He then began to think of breaking *S* into small strips and used a *representative strip* to structure his understanding.

Brian: I believe that, uh, I'm just trying to relate this to rectangles. If we just took the area of this piece of the rectangle here, this part of the table, and found the total force exerted on that, you would get some kind of estimate… [Draws a large rectangle to represent the table.] Let's just say this is *dA* [references a small strip at one end of the table]. This whole thing [i.e. the strip] is *dA*… And you have pressure pushing on that, on all that area. So you can multiply *P* times *dA* and you get the total force pushed, exerted on that part of the table… Yeah, if you make that area smaller and smaller and smaller and then add up those infinite, those really small areas on the whole table, you get the total force.

By drawing on the *adding up pieces* symbolic form, both Curtis and Brian were able to push past the difficulties they encountered by relying on anti-derivatives and areas under curves. This notion gave them a conceptual framework for understanding how the integrand and differential interacted within each small piece to give a small amount of the resultant quantity. All of the small amounts of force were then added up to capture the *total amount* of force. Their language suggests they were both more confident and satisfied by their *adding up pieces* explanations. Thus, this conceptualization appears to be *highly productive* for understanding physics integrals.

Conclusion

There are several ways to interpret the definite integral, including notions of areas, values of anti-derivatives, and summations, each of which appears useful for simple, basic mathematics integrals. The results of this study, however, show that in the physics context the *adding up pieces* symbolic form proved most productive for making sense of integrals. This is *not* to say that areas under curves and anti-derivatives are not important or should not be taught. Rather the results suggest that understanding the integral as an addition over many pieces is a key idea that should receive at least an equal share of attention in calculus courses. This corroborates with conclusions by others that the idea of *accumulation* is particularly helpful for fully understanding integration (Thompson, 1994; Thompson & Silverman, 2008). The results demonstrate a need for curriculum and instruction to develop this particularly important way to interpret the integral, especially in order to support the application of this important concept to science.

References

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics, 62*(8), 750-762.

Bressoud, D. M. (2012). MAA calculus study: Intended careers. *Launchings: Mathematical Association of America.* Retrieved December 13, 2012, from http://www.maa.org/columns/launchings/launchings_10_12.html.

Christensen, W. M., & Thompson, J. R. (2010). Investigating student understanding of physics concepts and the underlying calculus concepts in thermodynamics. *Proceedings of the 13th special interest group of the Mathematical Association of America on research in undergraduate mathematics education*. Raleigh, NC.

Engelke, N., & Sealey, V. (2009). The great gorilla jump: A Riemann sum investigation. *Proceedings of the 12th special interest group of the Mathematical Association of America on research in undergraduate mathematics education*. Raleigh, NC.

Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York, NY: Basic Books.

Hall, W. L. (2010). Student misconceptions of the language of calculus: Definite and indefinite integrals. *Proceedings of the 13th special interest group of the Mathematical Association of America on research in undergraduate mathematics education*. Raleigh, NC.

Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics, 68*(S1), S52-S59.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169-190). Mahwah, NJ: Erlbaum.

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89-120). Greenwich, CT: Information Age Publishing.

Hibbeler, R. C. (2012). *Engineering mechanics: Statics* (13th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Jones, S. R. (2013). Understanding the integral: Students' symbolic forms. *The Journal of Mathematical Behavior, 32*(2), 122-141.

MacLachlan, G., & Reid, I. (1994). *Framing and interpretation*. Melbourne, Australia: Melbourne University Press.

Mustoe, L. (2002). Mathematics in engineering education. *European Journal of Engineering Education, 27*(3), 237-240.

Pollock, E. B., Thompson, J. R., & Mountcastle, D. B. (2007). Student understanding of the physics and mathematics of process variables in PV diagrams. *AIP Conference Proceedings, 951*, 168-171.

Sealey, V. (2006). Definite integrals, Riemann sums, and area under a curve: What is necessary and sufficient? *Proceedings of the 28th annual meeting of the North American chapter of the International Group for the Psychology of Mathematics Education*. Merida, Mexico.

Sealey, V., & Oehrtman, M. (2005). Student understanding of accumulation and Riemann sums. In G. Lloyd, M. Wilson, J. Wilkins & S. Behm (Eds.), *Proceedings of the 27th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Eugene, OR: All Academic.

Serway, R. A., & Jewett, J. W. (2008). *Physics for scientists and engineers* (7th ed.). Belmont, CA: Thomson Learning.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1-36.

Sherin, B. L. (2001). How students understand physics equations. *Cognition and Instruction, 19*(4), 479-541.

Stewart, J. (2012). *Calculus: Early transcendentals* (7th ed.). Belmont, CA: Brooks/Cole.

Strauss, A., & Corbin, J. M. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Thomas, G. B., Weir, M. D., & Hass, J. (2009). *Thomas' calculus* (12th ed.). Boston, MA: Pearson.

Thompson, P. W. (1994). Images of rate and operational understanding of the Fundamental Theorem of Calculus. *Educational Studies in Mathematics, 26*(2-3), 229-274.

Thompson, P. W., & Silverman, J. (2008). The concept of accumulation in calculus. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp. 117-131). Washington, DC: Mathematical Association of America.

STUDENT VIEWS ABOUT TRUTH IN AXIOMATIC MATHEMATICS

Brian Katz
Augustana College

An undergraduate mathematics major should come to hold appropriate views about the conclusions reached by our disciplinary methods. This project explores the views about truth in axiomatic mathematics of a group of students who are (mostly) in their final proof-based course, Modern Geometry. Do these students hold expert-like views about truth in mathematics, and do those views change during a course that emphasizes epistemological themes? I find preliminarily that many of these experienced students do not distinguish the truth-value of theorems from that of definitions or axioms at the start of the term, but they develop more expert-like perspectives on truth during the course.

Key words: epistemology, truth, axiomatic mathematics, Geometry, concept maps

Background and Research Questions

One goal of an undergraduate program is that its graduates internalize an epistemic stance toward a discipline that approximates the stance of an expert and with which the graduates can evaluate and contextualize claims made in that discipline; this general goal can be specified to each discipline and at different levels. One such specification in mathematics is captured in the following articulated view: Once axioms are chosen for a domain within mathematics, they are viewed as true; statements in this domain are theorems if they can be proved rigorously starting from the axioms, and as a result these theorems are true as well.

At my institution, junior mathematics majors generally take three proof-based courses: Abstract Algebra then Real Analysis then Modern Geometry. The first two courses are required of all mathematics majors; Geometry is required only of the pre-service secondary education majors. Modern Geometry is usually the final proof-based course for these students and is the most explicitly axiomatic course in the department. Moreover, the course includes activities designed to help students focus on the nature of axiomatic mathematics and reflect on their own views about axiomatic mathematics. As a result, Modern Geometry offers an exciting position in the program at which to assess the views held by the students about truth in axiomatic mathematics, which brings us to my research questions.

1. Do these Modern Geometry students hold expert-like views about truth in mathematics, whether at the beginning or end of the course?
2. Do these Modern Geometry students' views about truth in mathematics change during the course; if they change, in what ways do they change?

These questions necessitate a cognitive approach to the data from this course; however, I hope to turn eventually to related questions about the complex social interactions that led to nuclear classroom episodes and the establishment of norms for rigor and truth.

In 1991, Sfard described a framework for integrating the operational and structural facets of mathematical conceptions that are "dual" and "complementary". These conceptions are conjectured to move from procedural to structural through phases that she calls *interiorization* (becoming familiar with the processes involving the new conception), *condensation* (squeezing lengthy sequences of operations into manageable units), and *reification* (conceiving of the conception as an integrated object). Sfard's examples of conceptions include "number" and "function"; college students are certainly expected to have reified these conceptions in some form by the end of their first year in college. In lower-division courses, the operational facets of a conception are generally computational; however, in upper-division courses, proof becomes one of predominant operational facets. Just as arithmetic eventually becomes reified and studied as an object in algebra, an extension of

Sfard's framework might predict that proof would become reified near the end of an undergraduate mathematics major and studied as axiomatic mathematics.

Other researchers have investigated ways in which students validate proofs of individual theorems. For example, Alcock and Weber (2005) investigate student line-by-line validation of proofs. Using the language of Toulmin's (1969) framework for proof validation, they give students a short proof in which all *data* and *conclusions* are valid but one implicit *warrant* is false. They find that Toulmin's framework explains the behavior of the majority of their students, though a large portion require prompting to consider the implicit warrant. In theory, viewing truth in an axiomatic system could be approached like validating a large argument, with the axioms serving as data, the theorems as conclusions, and the proofs as warrants. In this manner, Toulmin gives us a few critical terms and connections to observe.

As a clarifying example, let's try to use the framework from Sfard to interpret Alcock and Weber's findings. The students who (correctly) see the prompting proof as invalid notice the false implicit warrant, which is evidence of proof (or proof validation) having been interiorized; these students may have condensed proof validation as a conception, and the difference between those who were successful with and without prompting might be the thoroughness or appropriateness of that condensation. The students who reject the proof because "it doesn't use definitions" may have a naïve or partial interiorization. Those students who accept the proof as valid don't seem to think of proof validation as a process and hence may be struggling to start interiorizing proof (at least in this content domain).

Using this extended version of Sfard's framework, I am interested in understanding the extent to which proof has been reified as a conception for my students and how the structural aspects of this conception compare between the students and an expert. Importantly, Sfard concludes that the development of conceptions will be hierarchical and that reification at a lower level and interiorization at the next higher level are "prerequisites for each other". Her prediction suggests that my Modern Geometry students may reify proof only as they interiorize the axiomatic method during the course, not before.

Data and Methods

The students in Modern Geometry engaged in a concept-mapping task centered on "mathematical truth" that included a detailed explanation of concept-mapping and a list of 15 other required terms such as "proof", "axiom", "rigor", and "definition". The students engaged with exactly the same task on the first and last days of the course, and their maps were not discussed explicitly between these events. After the second mapping event, the students were given their original concept-map and asked to compare and contrast their two concept-maps in their own words. I also collected multiple reflective writings produced by the students in response to prompts and (sometimes) course readings. The data for this project represent artifacts from 19 students from two sections of the course, taught quite similarly by the same instructor (the author) over two consecutive years.

A mathematician (the author) processed all of the concept maps by first coding each connection as accurate/appropriate, inaccurate/inappropriate, or vague/confusing. I then looked for terms that appeared as endpoints in inaccurate or confusing connections among multiple students; I also compared the connections made by the same student at different times and then read the student's own interpretation of those differences. Similarly, I scanned passages from the student reflective writings for epistemological themes and coded them as accurate, inaccurate, or confusing.

Preliminary Results

It may not surprise you, given the abstract nature of the concept-mapping task, that the pre-course maps contain large numbers of confusing connections and small numbers of both accurate and inaccurate connections. Happily, the post-course maps contain large numbers of accurate connections and small numbers of inaccurate and confusing connections. Somewhat surprisingly given the level of confusion, there were some patterns to the inaccurate connections in the pre-course maps. In particular, most of the students put either “definitions” or “axioms” (or both) as “sister” connections to “theorems/corollaries” or “lemmas”. This indicates to the author that most of the students did not or could not differentiate the truth-value of a definition (which do not have truth-values) from that of axioms (which may be seen as choices for “true”) and theorems/corollaries (which should be seen as consequences hence “true”) at the start of the term.

The concept-maps demonstrate at least three related approaches to framing mathematical truth, generally observed through the implied subject of the verbs used to label connections in the map. In a *psychological* map, the student explains a way that an individual mind comes to *learn* about the truth of certain mathematical conceptions. In a *procedural* map, the student explains a way that a theorem comes to be proved, focusing on the *process* of doing mathematics. In an *epistemological* map, the student explains a way that *truth flows* through an axiomatic system, from choices to conclusions.

The Appendix contains an example pair of pre- and post-course maps from the same student. The pre-course map is an example of a map that does not separate “axioms” from “theorems/corollaries”; this map is mostly of the epistemological type. The post-course map is impressively accurate/appropriate and highly procedural.

The student writings provide a more nuanced window into each student’s views. One interesting theme that has bubbled to the surface addresses the truth-value of axioms. From my perspective, it is equally reasonable to think of axioms as true or as choices (and hence all other truth is relative to those choices). Most of the students found Kant’s notion of a “synthetic a priori truth” extremely comforting because it allowed them to return to a position from which the axioms are true in some absolute sense and hence mathematical truth is a form of universal truth. The rest of the students were more comfortable accepting the truth of mathematical statements as contextual, with choices of axioms serving as part of the world in which mathematics is done; from this position, truth is relative and yet somehow not random or arbitrary. The number of students is small, but the pre-service teachers almost all accepted axioms as absolutely true while the “pure” mathematics majors generally accepted axioms as the foundation for relative truth.

Using the language derived from Sfard above, each student appears to have moved from interiorizing or condensing proof into at least partial reification by the end of the course. Without this reification, the concept-mapping task would be quite difficult to parse, a claim that might explain the large number of vague/confusing connections in the pre-course maps.

Discussion

I have several concerns about the methods and analysis above and am sure there are other, possibly serious concerns. (i) The student concept-maps are analyzed mostly by comparison to one, un-articulated map in the mind of the author; not only does this put reliability into question, it tacitly assumes that my map is an accurate representation of an imaginary map held by all mathematicians or a universal map outside of practitioners. (ii) It is not clear how appropriate it is to use a concept-mapping task this abstract to measure the student views about truth. As mentioned above, the pre-maps could largely be the result of the students’ inability to parse the task. At the very least, Alcock and Weber demonstrate an effect from prompting, which this task does by its very design. (iii) More generally, as a mathematician who is new to education research, I do not feel at all secure in my ability to use and articulate appropriate theoretical frameworks and methodologies. I am quite certain that my own worldview is at play and almost unchecked in the analysis above. Some questions:

1. Are there frameworks or methodologies that I should employ to add rigor to my analysis? More generally, is there any reason to believe that I am accessing the same aspect of student thinking as any other extant approach?
2. Are there modifications of my concept-mapping task that would make it less likely to distort the form of my students' views? Are there other types of data I should use to triangulate the results of my analysis?
3. What are specific questions that I might ask about the classroom environment in light of the data about student cognition above?

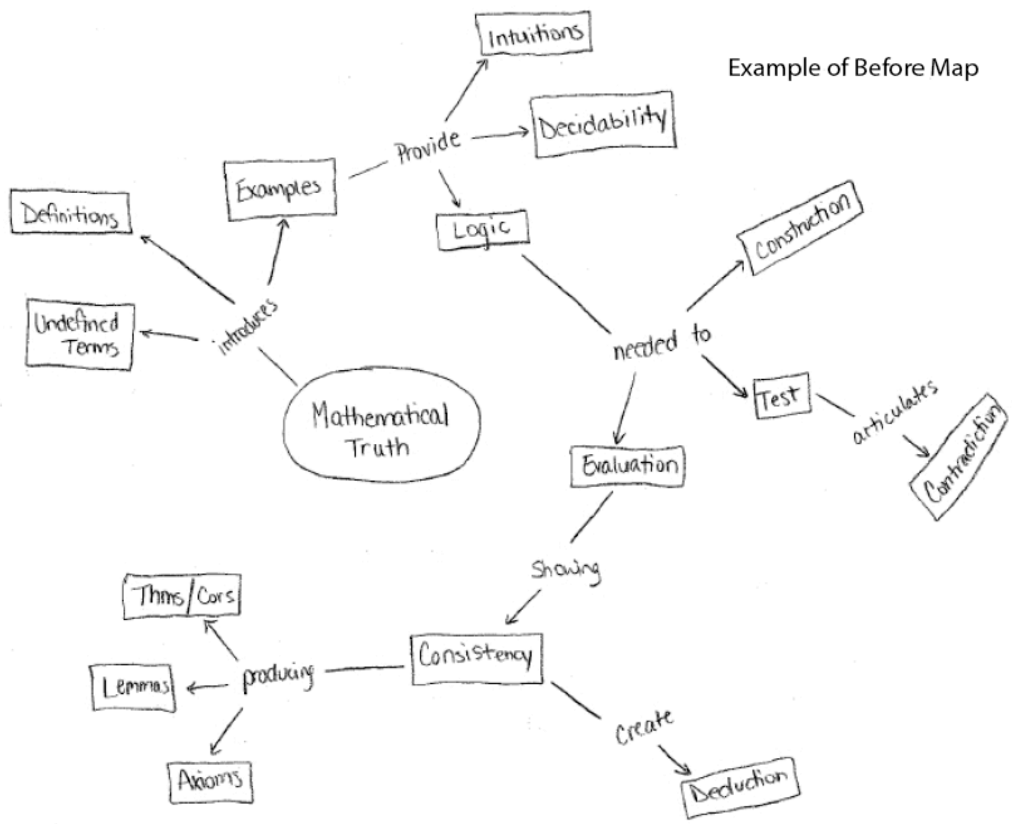
References

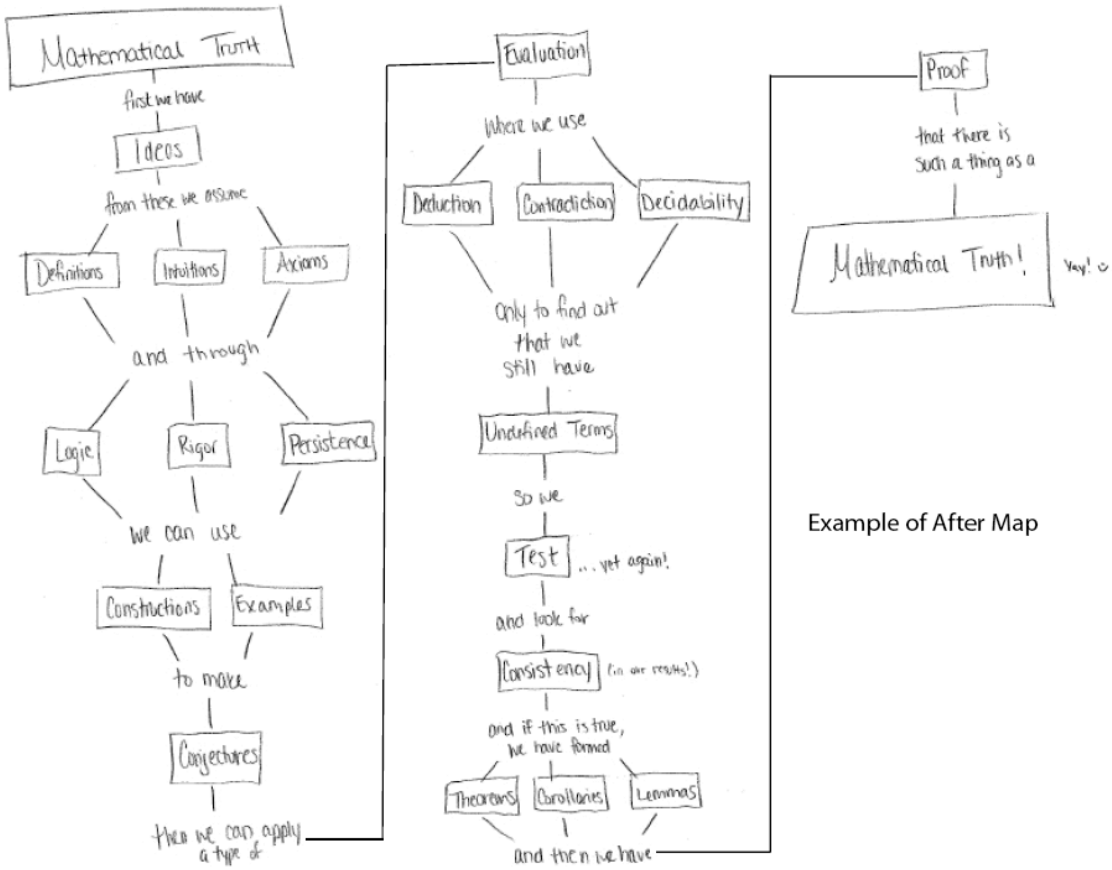
Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior, 24*, 125-134.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects on different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

Toulmin, S. (1969). *The uses of argument*. Cambridge: Cambridge University Press.

Appendix





Example of After Map

77

TEACHING INQUIRY-BASED MATHEMATICS TO IN-SERVICE TEACHERS: RESULTS FROM THE FIELD

Karen Allen Keene
North Carolina State University

Celethia McNeil
North Carolina State University

We present results from a classroom teaching data collection that involved practicing teachers as they participated in an inquiry-oriented differential equations (IO-DE) course. Data was collected to investigate how the teachers' participation in this kind of course, different from any of their previous mathematics courses, may influence their conceptions of teaching, mathematics, and student learning. Preliminary results indicate that the perceptions of teachers were changed by their experience in the class, at least as expressed in interviews. The teachers were likely to attempt to use more student-centered methods in their classrooms and believe that student learning is better in the student-centered environment. Additionally, attitudes about non-lecture, although mixed, did indicate a positive tone towards the constructivist perspective on learning. Finally, the teachers' participation in argumentation increased during the IO-DE course.

Keywords: inquiry, teacher change, differential equations, teacher conceptions

The President's Council of Advisors on Science and Technology (PCAST) recommended in their recent report that the United States needs to improve a) training for Science, Technology, Engineering, and Mathematics (STEM) teachers and b) undergraduate STEM education (2012). Specifically, the call has gone out to increase the quality of undergraduate STEM education, integrate the different STEM subjects in interdisciplinary ways, and increase the quality and number of K-12 STEM teachers. In this preliminary report, we report on one new direction we are taking to address these issues by studying an undergraduate level student-centered mathematics curriculum that was implemented at the master's level for high school mathematics teachers.

In this research, we used previously developed inquiry-oriented differential equations (IO-DE) materials that had been taught at the undergraduate level and taught it in a graduate mathematics course for teachers. We investigated to see if there is evidence that this type of instruction changes attitudes and orientations of teachers, and affects instruction when the teachers go back to the classroom. We also investigated whether there is an increase in understanding of mathematics, teachers' willingness to discuss mathematics, and their motivation to use more contextual based teaching.

Literature Review

Part 1. IO-DE research

Inquiry-oriented differential equations (IO-DE) (Rasmussen & Kwon, 2007) has been developed and used in a number of classrooms in the United States and internationally. Several research publications have provided evidence that students' participation in this student-centered course allows them to develop a conceptual understanding of solutions to differential equations (Rasmussen, Kwon, Allen, Marrongelle, & Burtch, 2006), retain the knowledge better (Kwon, Rasmussen, & Allen, 2005) and find ways to use their prior knowledge to understand differential equations (Author, 2008). The course focuses on differential equations as dynamic rate of change equations and emphasizes autonomous differential equations, and systems of differential equations. The students work in cycles of small group and whole class.

The research base developed in IO-DE includes results that elaborate the construct of students' dynamic reasoning (Author, 2007) as well as publications that discuss results including, but not exhaustive, emergent models (Rasmussen & Blumenfeld, 2007), technology's role in learning undergraduate mathematics (Keene & Rasmussen, 2013) and building on realistic starting points to help students construct understanding of the mathematics (Rasmussen, 2001).

Part 2. Research on what influences teachers' instruction

Space does not permit discussion of the large body of literature that study teachers' practices and the influence on such practices. Stein, Remillard, and Smith (2007) provide a nice overview and use the notion of "enacted curriculum". Early research in this area has led researchers to hypothesize that one true statement about teachers is "Teachers teach how they were taught" (see Zeichner & Tabachnick, 1981 for an example). There have been many professional development workshops and other experiences to provide ways for teachers to experience the learning of mathematics in a more constructivist environment, but as professional development Sztajn, 2011). There is less research about how teachers' participation in university classes after they have been teaching might make a difference (see work from Arizona State University for example). We use the word "conception" as Lloyd and Wilson (1998) used it: "to refer to a person's general mental structures that encompass knowledge, beliefs, understandings, preferences, and views" (p. 249) and meld this with the enacted curriculum description for our theoretical framework. To that end, we ask the following questions:

- How do practicing teachers in a student-centered post-calculus math course change their conceptions of mathematics, teaching and learning, and the value and usefulness of mathematics?
- How does it affect teachers' practice?
- How does the significant use of argumentation in mathematics look in an advanced mathematics course and how does it improve student understanding and views of mathematics?

Research Methods

Setting and Participants

The course where the research took place was a master's level mathematics course for high school mathematics teachers pursuing their master's degree in a large southeastern university. The instructor was also the primary researcher and had experience teaching the specific course. There were 20 students in the course, 14 of whom were enrolled in the mathematics education masters program and teaching in high schools or community college. These 14 were the participants in our study. Their teaching experience ranged from 0 to 6 years. Additionally, there were six PhD students who were enrolled in the course and were participants in the study. They did however play a significant role in the class discussions and contributed to the student-centered environment.

Data Collection

Following is the data collect that are relevant to the proposed presentation:

- Differential Equations Relational Understanding Assessment (Author, 2011). A quantitative instrument used to determine how conceptual understanding of differential equations improved.
- Class Video-Taping. The class met 14 times for 2 hours and 45 minutes for each class session. Two cameras were used to tape all of the whole class discussions, where one camera focused on the teacher and the other on the students. During small group work, the cameras were focused on one of the small groups to understand what transpires in that setting.
- Interviews. Approximately 9nine months after the end of the class, we conducted one hour interviews with 8 of the participants in the class. These interviews were semi-structured and led by one of the authors. The questions were designed to answer the research questions and participants were allowed to elaborate on answers as they desired.

Data Analysis

We are in the early stages of analyzing the data and report on some of the analysis at this time. Specifically, we have been working on the data from the surveys and final exams, whole class discussions and the individual interviews. We first watched all the class videos and broke the class discussions into segments. Each segment is 15 minutes focused on one particular mathematical task, presentation, or dialogue. We then transcribed the videos while noting important discourse and any interesting observations from the video.

The interviews were conducted and transcribed. We used axial coding (Strauss & Corbin, 1998) to develop a set of codes for each of the statements made by the interviewed teachers. One of the researchers did the primary coding and the other researcher did confirmatory checking. The codes were either identified as pedagogical or mathematical and the codes were refined and combined where appropriate. Plans are for themes to emerge; we will report on these themes at the conference.

Results

Some of the primary codes that we think will tie to the important themes are listed in Table 1. We provide a brief discussion and illustration of two themes here and will discuss all the themes in the presentation.

Table 1.

Theme	**Mathematical or Pedagogical**
Frustration in different types of instruction	**Pedagogical**
Want to make changes to teaching in own classroom	**Pedagogical**
Tied together earlier mathematical knowledge	**Mathematical**
Context was useful and important to learning	**Mathematical/Pedagogical**
Differential equations as rate of change was important	**Mathematical**

Preliminary results indicate that teachers have made changes in beliefs about learning in mathematics, which applies to the first research question. Six of the teachers made statements in their interviews that indicate that using this kind of instruction was something they would like to implement in their classroom. For example,

Interviewer: Did your experience in this class make any difference in the way you teach? Why or why not?

Teacher 1: Yea, um...I...as much as I hated that type of student-centered... when I did figure something out, I was really pumped and I got really excited about it. And so every unit or two, I'd threw something in where they had to figure out why is this the formula...or what does this constant really mean in this equation and that kind of stuff.

Mathematically, teachers remembered some of the content after nine months, specifically, they recalled some of the more difficult concepts such as analyzing phase plans and how they remembered it because of the kind of instruction.

> Teacher 2: Or I really didn't have a good understanding of it or something, but by the end of that class, I really understood the idea and I understood how it relates to differential equations. So, for me it was big, that part.

Results for Research Questions 2 and 3 will be discussed at the presentation.

Conclusion and Implications for Teaching

These results, although found in a single classroom, provide some ideas of what the future use of more student-centered instruction in preservice and inservice teacher programs could look like and contribute. We believe that this applies to both practicing teachers and preservice teachers that are still in an undergraduate mathematics program. It is not enough to teach our future mathematics teachers in a more student-centered class in their methods classes, but to also influence those who are currently practicing. Additionally, we need to introduce more mathematics classes to this style of instruction. In the long term, this may will more significantly affect mathematics education at the K-12 level.

Questions for Audience

1. Should mathematics instruction for preservice teachers be different than for mathematics majors? How and why?
2. How important is context when teaching differential equations and/or other undergraduate mathematics courses?
3. Do you think this line of inquiry is useful and relevant? How could we extend it into the undergraduate classroom?

References

Keene, K. & Rasmussen, C. (2013). Sometimes less is more: Examples of student-centered technology as boundary objects in differential equations. In S. Habre (Ed.). *Dynamical mathematical software and visualization in the learning of mathematics* (33 pages). IGI Global: Hershey, PA.

Kwon, O. N., Rasmussen, C., & Allen, K. (2005). Students' retention of knowledge and skills in differential equations. *School Science and Mathematics*, *105*(5), 227-239.

Lloyd, G. M., & Wilson, M. (1998). Supporting innovation: The impact of a teacher's conceptions of functions on his implementation of a reform curriculum. *Journal for Research in Mathematics Education,* 29(3), 248–274.

President's Council of Advisor's on Science and Technology. (2012). Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. *Report to the President.* Washington, DC: The White House.

Rasmussen, C. (2001). New directions in differential equations: A framework for interpreting students' understandings and difficulties. *Journal of Mathematical Behavior*, 20, 55-87.

Rasmussen, C., & Blumenfeld, H. (2007). Reinventing solutions to systems of linear differential equations: A case of emergent models involving analytic expressions. *Journal of Mathematical Behavior,* 26, 195-210.

Rasmussen, C., & Kwon, O. (2007). An inquiry oriented approach to undergraduate mathematics. Journal of Mathematical Behavior, 26, 189-194.

Rasmussen, C., Kwon, O., Allen, K., Marrongelle, K. & Burtch, M. (2006). Capitalizing on advances in K-12 mathematics education in undergraduate mathematics: An Inquiry-oriented approach to differential equations. *Asia Pacific Education Review*.Vol. 7, No. 1, 85-93.

Stein, M., Remillard, J., & Smith, M. (2007). How curriculum influences student learning. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 1, pp. 319–369). Charlotte, NC: Information Age Publishing.

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Thousand Oaks, CA: Sage.

Sztajn, P. (2011). Standards for reporting mathematics professional development in research studies. *Journal for Research in Mathematics Education*, *42*(3), 220-236.

Zeichner, K., & Tabachnick, B. R. (1981). Are the effects of university teacher education washed out by school experience? *Journal of Teacher Education*, 32, 7-11.

THE EFFECTIVNESS OF 5-MINUTE PREVIEW VIDEO LECTURES USING SMART BOARD, CAMTASIA STUDIO, AND PODCASTING ON MATHEMATICAL ACHIEVEMENT AND MATHEMATICS SELF-EFFICACY

Ph.D. Minsu Kim
University of North Georgia

The purpose of this study is to examine the effectiveness of 5-minute preview video lectures for each lecture using podcasting in terms of mathematical achievement and mathematics self-efficacy in intermediate algebra and college algebra courses at a university. Data from 128 students in 6 sections collected for two semesters through first and final exams, questionnaires, a classroom observation checklist, and the Mathematics Self-Efficacy Scale. The preliminary findings indicate no significant difference on the mathematical achievement and mathematics self-efficacy between the control group who did not watch the preview lectures and the treatment group who watched the preview lectures while the treatment group slightly developed their mathematics self-efficacy and abilities for mobile technology. In addition, the treatment group was significantly satisfied with the preview lectures. When the treatment group was divided into intermediate low and high subgroups based on the first exam, the intermediate low subgroup significantly improved their mathematical achievement.

Key words*:* Preview Video Lectures, Podcasting, Mathematical Achievement, Mathematics Self-Efficacy, Mobile Technology

Introduction

Mathematics instructors at colleges and universities have improved pedagogical environment both inside and outside the classroom for students through technology. Because of the technologies such as class websites, educational software, Smart Boards, and podcasting, mathematics instructors also have developed their resources in terms of how content is delivered after their lectures. For example, students are able to find electronic versions of PowerPoint slides or PDF handouts on the class websites (Copley, 2007). After the appears of podcasting, instructors were able to record their lectures as streaming videos and distribute them through podcasting in order to help students learn materials at their convenience (Laing et al.,2006; Sharples, 2000). Even though instructors provide the opportunities for students to learn the materials on class websites, a number of students attend class with copies of the lecture notes or handouts without reading the materials before their classes. How do mathematics instructors encourage students to be interested in up-coming lectures through the technologies?

While podcasting has become a popular media on the Internet (Searls, 2005), researchers in higher education have been interested in podcasting as an educational tool. Duke University evaluated the effectiveness of mobile learning such as iPods from more than 1600 students in learning and teaching in 2004 (Belanger, 2005). According the Belanger’s research in 2007, students’ achievement had improved in quality at the same time as the students’ motivation and use of resources online had increased (Sutton-Brady, Scott, Taylor, Carabetta, & Clark, 2009). Several studies have shown that streaming video significantly influences students’ achievement in higher education (Carlson, K., 2009; Mark, S., 2004; McGrann, R. T. R., 2005; Reed, R., 2003). In addition, instructors have used podcasting in order to provide podcast streaming video recordings of lectures for students to review and revise after their classes (Laing et al., 2006).

Research has also provided several advantages for students’ learning in terms of use of podcasting in higher education. Students are able to revise and study content through replaying podcast episodes on diverse devices (Laing, Wootton, & Irons, 2006; Shannon, 2006). In addition, students have opportunities to manage their times in order to learn content instead of only scheduled lecture times (Sharples, 2000). Other researchers have studied supplementary podcasts, which are short podcasts including 5-minute summaries of presentations (Calder, 2006), interviews for past and upcoming lecture content, and announcements (Bell, Cockburn, & Wingkvisit, 2007). Even though researchers have raised their concerns for applications of podcasting as an innovative tool in education, there is still limited evidence on the pedagogical strategies for student learning in mathematics education and the effectiveness of podcast video lectures in terms of students' mathematical achievement and mathematics self-efficacy. In addition, there are few studies on the effect of short-format podcasting of upcoming core-lecture content regarding students' performance and their mathematical self-efficacy.

The purpose of this study is to examine the effectiveness of preview video lectures about 5 minutes in length for each lecture using Smart Board, Camtasia Studios, and Podcasting in terms of mathematical achievement and mathematics self-efficacy in intermediate algebra and college algebra at a university. The aim of this research is to answer the following two research questions and two sub-questions: 1) How do students who watch preview lectures and students who do not compare in their mathematical achievement? Is there a significant difference in mathematical achievement between two groups in terms of the average final examination scores? 2) How do students who watch preview lectures and students who do not compare in their mathematics self-efficacy? Is there a significant difference in mathematics self-efficacy between the two groups? Two hypotheses were tested in this study: 1) The mean of final exam of students who watch the preview video lectures is not significantly higher than the mean of final exam of students who do not watch the preview video lectures. 2) The mean of mathematics self-efficacy scores of students who watch preview lectures is not significantly higher than the mean of mathematics self-efficacy scores of students who do not watch preview lectures.

Theoretical Rationale

Increasing the use of portable devices, podcast video lectures has become a potential tool for learning and teaching in higher education regarding flexibility (Kukulska-Hulme, Traxler, and Pettit, 2007; Traxler, 2008). In addition, Chan and Lee (2005), Chan, Lee, and McLoughlin (2006), Bell, Cockburn, and Wingkvist (2007), and Sutton-Brady, Scott, Taylor, Carabett, and Clark (2009) suggested short-format podcasts have more benefits for student learning rather than an hour-long recorded podcasts have regarding the pedagogical design of podcasts. Based on these studies on effectiveness of podcasting, students will have the opportunities to improve their readiness for the next classes and encourage students to be interested in lectures, utilizing 5-minute podcast video lectures through their mobile devices. Finally, the use of 5-minute preview podcast video lectures will support students in mathematical achievement and mathematics self-efficacy. This study will show the use and effectiveness of short-format podcast video lectures in student learning and teaching in mathematics.

Methods

This study used a quasi-experimental pretest/posttest design to answer the research questions. The participants were 128 students in 6 sections, 3 intermediate algebra and 3 college algebra courses with the same instructor from the fall of 2012 to the spring of 2013 at a university. The schedules of the 6 sections were between 9:00 am to 5:00 pm, not evening classes. The participants enrolled in the classes according to their schedule and other preferences. In this study, 59 students participated, consisting of 39 in 2 intermediate algebra courses and 20 in 1 college algebra course in the fall of 2012. There were 38 male and 21 female students in the fall of 2012. The participants served as the control group and did not watch the preview video lectures. In the spring of 2013, 69 participants served as the treatment group and watched the preview video lectures before attending each class. The 69 participants consisted of 19 in 1 intermediate algebra course and 50 in 2 college algebra courses. In addition, there were 24 male and 45 female students in the sections of spring 2013. If students in intermediate algebra courses participated in this study in the fall of 2012, the students were removed from the study of the spring 2013 classes. I employed three different data sources: the first and final exams, questionnaires, a classroom observation checklist, and the Mathematics Self-Efficacy Scale (Betz & Hackett, 1993). The exams were the same question formats with different numerical values between the two semesters.

I collected the data of the first exam as a pretest and the final exam as a posttest each semester in order to measure the students' mathematical achievement. After the first exam, the participants completed the first questionnaire with a consent form and the pre-Mathematics Self-Efficacy Scale. The first questionnaire consisted of background information and abilities in mobile technology. In addition, I gathered the second questionnaire, satisfaction of the 5-minute preview video lectures, and the post-Mathematics Self-Efficacy Scale at the end of semester.

To answer the research questions, I performed t-tests and ANOVAs to compare the mathematical achievement and the Mathematics Self-Efficacy between the control and treatment groups, using Minitab 16, software for statistics. The independent variables were class type: non-preview video lectures (NPL) and preview video lectures (PL); the first exam score, in order to measure incoming level of mathematics: intermediate low (IL) and high (IH) groups; gender; and class attempts: first time student (FS) and repeating student (RS). The dependent variables were mathematical achievement by the final exam and the post-mathematics self-efficacy. For the first and second null hypothesizes, I conducted two independent-sample t-tests. Similarly to Cohen (2001), I used a significance level of 0.05 for this study. In addition, I performed ANOVA to compare the pretest to posttest changes in mathematical achievement and the mathematics self-efficacy scores of the control and treatment groups.

Preliminary Findings and Discussion

By the two independent-sample t-tests, there was no significant difference on the mean of first exam scores between the control and treatment groups, even though the mean of first exam scores of the control group was higher than the mean of first exam scores of the treatment group. In addition, the treatment group slightly developed their mathematics self-efficacy and abilities for mobile technology, although there was no significant difference on the mathematical achievement and mathematics self-efficacy between the control group who did not watch the preview lectures and the treatment group who watched the preview lectures. Moreover, the treatment group was significantly satisfied with the preview video lectures by an independent-sample t-test. When the treatment group was divided into intermediate low and high subgroups based on the first exam scores, the intermediate low subgroup significantly improved their mathematical achievement because $p=0.028$ is less than $\alpha=0.05$. The preliminary findings indicate that the preview video lectures significantly influence the lower-intermediate students' mathematical achievement.

Even though there are several limitations in terms of the number of students and the different courses and semesters, the study will show the effectiveness and pedagogical design of short-format podcast video lectures. The preliminary results of this study will have potential implications for mathematics faculty and online system developers of textbooks and contribute to our knowledge of pedagogical approaches outside the classroom, using technologies in mathematics education.

Questions to the Audience

What theoretical rationale is appropriate for this study?
What other methodologies could I use to analyze the data?

References

Belanger, Y. (2005). *Duke University iPod first year experience final evaluation report.* Available online at: at: http://cit.duke.edu/pdf/ipod_initiative_04_05.pdf (accessed 8 March 2007).

Bell, T., Cockburn, A., Wingkvist, A., & Green, R. (2007). Podcasts as a supplement in tertiary education: an experiment with two computer science courses. In *Paper presented at the mobile learning technologies and applications conference*. (Auckland, New Zealand).

Betz, N. E., & Hackett, G. (1993). *Mathematics self-efficacy scale sampler set: Manual, test booklet and scoring guide.* Palo Alto, CA: Mind Garden.

Calder, A. (2006). If you can't beat 'em, podcast 'em. Paper presented at the 9th Pacific Rim–First Year in Higher Education Conference, 29 June–1 July, Townsville, Queensland University of Technology. http://www.fyhe.qut.edu.au/past_papers/2006/Nuts%20&%20Bolts/Calder.pdf.

Carlson, K. (2009). Delivering information to students 24/7 with Camtasia. *Information Technology & Libraries*, 28(3), 154-156.

Cohen, B. H. (2001). Explaining psychological statistics (2nd ed.) New York: Wiley.

Copley, J. (2007). Audio and video podcasts of lectures for campus-based students: production and evaluation of student use. *Innovations in Education and Teaching International*, 44(4), 387-399.

Laing, C., Wootton, A. & Irons, A. (2006). iPod! uLearn?, *Current Developments in Technology-assisted Education.* Available online at: http://podcasting.thefutureoflearning.googlepages.com/514-518.pdf (accessed 8 March 2007).

McGrann, R. T. R. (2005). Enhancing engineering computer-aided design education using lectures recorded on the PC. *Journal of Educational Technology Systems*, 34(2), 165-175.

Mark, S. R. (2004). Personalizing the online classroom using Tech-Smith's Camtasia or Microsoft's Windows Media Encoder. *Online Classroom*, 4-5.

Patrick, H., & Hicks, L. (1997). Relations of perceived social efficacy and social goal pursuit. *Journal of Early Adolescence*, 17, 109-129.

Shannon, S.J. (2006). Why don't students attend lectures and what can be done about it through using iPod nanos? In *Who's learning? Whose technology? Proceedings of the 23rd annual ASCILITE conference*, ed. L. Markhauskaite, P. Goodyear, and P. Reimann,753–756. Sydney, NSW: University of Sydney.

Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. *Computers & Education* 34, nos. 3–4: 177–93.

Searls, D. (2005) DIY radio with PODcasting. Available online at: at: http://www.itgarage.com/node/462 (accessed 8 March 2007).

Sutton-Brady, C., Scott, K. M., Taylor, L., Carabetta, G., & Clark, S. (2009). The value of using short-format podcasts to enhance learning and teaching. ALT-J, *Research in Learning Technology*, 17(3), 219-232.

Reed, R. (2003). Streaming technology improves achievement: study shows the use of standards-based video content, powered by new internet technology application, increases student achievement. *T.H.E. Journal*, 30(7)

79

TRANSFER OF LEARNING: EXAMINING INDIVIDUALS IN SOCIAL SETTINGS

Jeffrey J. King
University of Northern Colorado
jeffrey.king@unco.edu

Stephenie Anderson
University of Northern Colorado
stephenie.anderson@unco.edu

Gulden Karakok
University of Northern Colorado
gulden.karakok@unco.edu

In this preliminary report, we share the design and results of the first phase of our on-going research study. Our three-phase study is designed to investigate individual student's transfer of learning of linear algebra concepts along with social mathematical interactions in which such concepts developed in group-based courses. We first frame our study in relation to current literature, then discuss our initial analysis from the first phase. Finally, we give a description of upcoming phases along with questions we wish to discuss with the audience.

Key words: Transfer of learning, Actor-oriented transfer, Linear algebra, Function

Many current transfer research frameworks adopt a cognitive lens to focus on an individual student's transfer of "knowledge pieces" or mathematical concepts from one task to another (see Mestre, 2005 for some of these approaches, more recent one Wagner, 2010, also some earlier studies e.g., Bassok, 1990; Gick & Holyoak, 1983). Some critiques of these frameworks highlight the need of sociocultural perspectives on transfer issues. This particular approach, however, becomes problematic as meanings of notions such as "learning", "understanding", "knowledge" create different constructs under such perspectives (Lave, 1988; Tuomi-Grohn & Engestrom, 2003; Hatano & Greeno, 1999). For example, within a cognitivist perspective, knowledge is generally viewed as a static object, acquired or constructed by an individual. In contrast, sociocultural perspectives assume a dynamic view of knowledge and its communal creation through communication. Thus, the definition of transfer itself becomes problematic under sociocultural perspectives. Under these critiques, new frameworks were developed to investigate issues further (e.g., Lobato, Rhodehamel, & Hohensee, 2012).

Seeing communication as an important aspect of learning and transfer is an important social addendum to our perspective and contributes to existing research (Lobato, Rhodehamel, & Hohensee, 2012). We approach transfer from a social constructivist perspective through implementation of the *Actor Oriented Transfer* (AOT) framework (Lobato, 2006), using contributions from Vygotsky's (1978) developmental processes in inter- and intramental planes to help us delineate actor's experiences. This allows us to focus on how an individual student sees a mathematical concept (function) as similar to another (transformation) in a target situation. In this case, we are not asking the question of "does a student transfer particular knowledge or mathematical practices?"; rather, we ask "what does a student transfer from their group-based linear algebra course to the target tasks, and in what ways do they see their learning experience as similar to the target situation (in regards to specific concept and practice)?" A careful reader might argue that we are interested in exploring if a student transfers function knowledge from the classroom to the interview tasks. However, this is not our goal and we believe transfer is not a static knowledge transfer. Given that, in any course, students are exposed to different mathematical concepts and practices, we need a narrower focus in our target tasks. Thus, our attention in data analysis will be given to the concepts of functions and linear transformations, and we will consider the mathematical practice (or strategy) of "making connections." For us, the "making connections" practice means exhibiting explicit connection-making efforts among experiences. In the linear algebra course, students will engage in this "making connections" practice by asking questions such as "Have I done anything similar before? Why do I think it is similar or different? If it is similar, what tools can I apply here? If not, what do I need to know?" These questions are purposefully context-free so that students can modify them to fit the activities in the classroom.

Our research questions for the main study are: 1) In what ways do students see interview tasks on linear algebra topics as similar to their prior experience in a group-based linear algebra course?; 2) In what ways do students make connections between the interview tasks and in-class practices?

Relevant Literature on Linear Algebra

Students' difficulty with transfer of learning in linear algebra has not been investigated enough (Karakok, 2009). Studies focusing on students' learning difficulties in linear algebra courses identified several issues, including: 1) Students' lack of ability to switch between different modes of thinking required in a linear algebra; 2) Students' lack of ability to connect different representations; and 3) Students' perception of axiomatic approach as pointless (Carlson et al., 1997; Dorier, 2000; Harel, 1989; Hillel & Sierpinska, 1994; Sierpinska, Nnadozie & Ortac, 2002; Stewart & Thomas, 2006). Results of such studies guided educators to develop instructional practices to help students (Siepinska, 2000; Wawro et al., 2012). Dubinksy (1997) highlighted that such attempts must consider students construction of their own ideas about important concepts. He further added that students must have a better understanding of background concepts that are not necessarily part of the course. He gave an "obvious example […] that having strong function concept is essential for understanding linear transformations" (p. 93). Therefore, we choose to focus on the concept of function and how students' conceptions of function relate to their conceptions of linear transformations.

Further research on the connection among functions and linear transformations was also investigated by Zandieh, Ellis, and Rasmussen (2012). Their research indicated three main resources students used as they worked on the interview questions: *properties, computations,* and *metaphors*. Researchers categorized students' statements as *property* if they indicated use of a property of a function or a linear transformation or a property of another related concept. *Computations* were used to indicate students' use of computation language or actions while working on tasks. *Metaphors* were used to identify different metaphors students called upon while reasoning. Researchers noticed that even though most of the participants agreed that linear transformations are a type of function, participants called upon different metaphors as they reasoned with linear transformations and functions. Such distinctions demonstrated by students are important in students' conception of linear transformations and transfer of learning of related linear algebra topics.

Methodology

Table 1 provides a summary of the phases of this on-going study. Our goal in the first phase of our study was to document students' pre-existing conception of function after taking the first semester calculus course but prior to taking an introductory linear algebra course. Results from the first phase will help us to understand how students could develop their understanding of linear transformation with their existing conception of function during the linear algebra course, and how they would use these experiences in the phase three interviews.

We interviewed one group of two students and another group of three students. These groups were selected from a first semester calculus course in which they worked together on labs. This course was video-taped throughout the semester to explore students' development of conception of function and to allow us to select groups of students who worked productively on in-class tasks.

Table 1 Summary of Phases of the Study

Phase 1	Students' prior conception of functions after taking Calc 1 (via group interviews)	Fall 2012, Spring 2013
Phase 2	Students' in-class group interaction (audio recording of groups and video recording of whole class)	Fall 2013
Phase 3	2 Individual-group-individual group interview cycles with 2 groups (up to 6 individuals)	Spring 2014

The data collected for Phase 1 consisted of four group interviews for each of the two groups, the first two interviews focusing mostly on functions and the second two interviews focusing on linear transformations. The goal of including questions on linear transformations was to investigate how the students would attempt to solve the problems using their existing conception of function.

Results

The preliminary interview data analysis indicated that participants had varying conceptions of function that appeared as they worked on the interview tasks. Four major themes surfaced: 1) vertical line test (VLT), 2) input/output relationships, 3) meaning of inverse, and 4) function descriptions.

At the outset, students were asked directly what they considered a function. They commonly answered with utterances such as a function being something that "passes the vertical line test." Upon further probing, students articulated their understanding of a relationship between input and output. For example, at the beginning of one interview, one participant stated

> *"… [the two graphs] do both pass the vertical line test because there's not two points right here , not two points right there, there's only one point. There's only one output for every input. So yeah they're both functions."*

Later in that same interview, the students stated,

> *"Ok [the equation] is a function. This is why I say it's a function. For every input you're gonna get a different output. There's no way I'm gonna get this output without this input. I can't choose a point right here and get this output."*

Contradictory statements such as these emerged from all participants in each group, and indicated a wavering conception of functions as input and output relationships.

When investigating linear transformations, two distinct meanings of inverse emerged from one group. First, students described an inverse function as one that "undoes" the original function. As they investigated the inverse of a given function, the students recalled the process of switching the x and y variables. From that point on, a second meaning of the word "inverse" emerged, used to describe any process or function that switched the x and y variables, even in two-variable functions.

The students used numerous descriptors of functions, including: formulas or equations, geometric transformations, a mapping from object to object, and a relationship that makes two things "go together." Our results provide similarities with the three resources described by Zandieh, Ellis and Rasmussen (2012). We plan to provide details on this aspect during our presentation.

Discussion

In our second phase, we will teach an introductory linear algebra course in which students will work on activities in their groups. As students work on new tasks each week, our instructional goal is to help them focus on strategies such as making connections. We plan to document any other emerging sociomathematical norms from this particular course. After this second phase, students who participated in the first phase and others will be invited to the third phase of the study. At our presentation, we hope to get feedback from the audience on the interview questions we have, and facilitate discussion on theoretical framing on transfer issues, specifically:

1) What are some mathematical practices that we should explore in the group and individuals interviews as we conduct interviews and analyze data?
2) What are some theoretical framing issues you notice from our paper and presentation, as well as in our on-going research?

References

Bassok, M. (1990). Transfer of domain-specific problem-solving procedures.*Journal of Experimental Psychology: Learning*, *16*(3), 522-533.

Carlson, D., Johnson, C.R., Lay, D.C., Porter, A.D., Watkins, A., & Watkins, W. (Eds.) (1997). *Resources for Teaching Linear Algebra,* The Mathematical Association of America Publication.

Dorier, J.L.(Ed.). (2000). *On the Teaching of Linear Algebra*, Kluwer Academic Publisher, Dordrecht.

Dubinsky, E. (1997). Some thoughts on a first course in linear algebra at the college level. In D. Carlson, C.R. Johnson, D.C. Lay, A. Duane Porter, A. Watkins and W. Watkins (Eds.), *Resources for Teaching Linear Algebra*, Mathematical Association of America, MAA, Notes Volume 42, pp. 85-105.

Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 16, 1-38.

Harel, G. (1989). Learning and teaching Linear Algebra: Difficulties and an alternative approach to visualizing concepts and processes. *Focus on Learning Problems in Mathematics, 11(2),* 139-148.

Hatano, H. & Greeno, J. G. (1999). Commentary: Alternative Perspectives on Transfer and Transfer Studies. *International Journal of Education Research, 31,* 645-654

Hillel, J., & Sierpinska, A. (1994). On one persistent mistake in linear algebra. In J.P. da Ponte & J.F. Matos (eds.), *Proceedings of the 18th International Conference for the Psychology of Mathematics Education, Vol. III,* 65-72. Lisbon, Portugal: University of Lisbon.

Karakok, G. (2009). *Students' transfer of learning of eigenvalues and eigenvectors: implementation of Actor-Oriented Transfer framework*. Doctoral dissertation, Oregon State University, 2009, Corvallis, OR.

Lave, J. (1988). *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. Cambridge, UK: Cambridge University Press.

Lobato, J. (2006). Alternative Perspective on the Transfer of Learning: History, Issues, and Challenges for Future Research. *The Journal of the Learning Sciences, 15(4),* 431-449.

Lobato, J., Rhodehamel, B., & Hohensee, C. (2012). "Noticing" as an alternative transfer of learning process. *Journal of the Learning Sciences*,*21*(3), 433-482.

Mestre, J. P. (2005). *Transfer of Learning: from a Modern Multidisciplinary Perpective*. Greenwich, Connecticut: Information Age Publishing.

Sierpinska, A. (2000). On some aspects of students' thinking in Linear Algebra. In J.L. Dorier (ed.) *On the Teaching of Linear Algebra* (p. 191-207), Kluwer Academic Publisher, Dordrecht.

Sierpinska, A., Nnadozie, A., & Oktaç, A. (2002). *A study of relationships between theoretical thinking and high achievement in linear algebra*. Concordia University. Unpublished Manusscript.

Stewart, S. & Thomas, M. O. J. (2006). Process-object difficulties in linear algebra: Eigenvalues and eigenvectors. *Proceedings of the 30th conference of the International Group for the Psychology of Mathematics Education*, Prague, Czech Republic.

Tuomo-Grohn, T., & Engestrom, Y. (2003). Conceptualizing transfer: From standard notions to developmental perspectives. In T. Tuomi-Grohn & Y. Engestrom , (Eds.), *Between school and work: New perspectives on transfer and boundary-crossing* (pp. 19-38). New York: Pergamon.

Vygotsky, L. (1987). The Development of Scientific Concepts in Childhood. In R. W. Rieber and A. S. Carton (eds.) *The Collected Works of L.S. Vygotsky*. Vol. 1, 167-241. (Original work published in 1934).

Wagner, J. F. (2010). A transfer-in-pieces consideration of the perception of structure in the transfer of learning. *the journal of the learning sciences*, *19*(4), 443-479.

Wawro, M., Rasmussen, C., Zandieh, M., Sweeney, G. F., & Larson, C. (2012). An Inquiry-Oriented Approach to Span and Linear Independence: The Case of the Magic Carpet Ride Sequence. *PRIMUS*, *22*(8), 577-599.

Zandieh, M., Ellis, J., and Rasmussen, C. (2012). Student concept images of function and linear transformation. Paper presented at the Fifteenth Conference on Research in Undergraduate Mathematics Education, Portland, OR.

Investigating Instructors' Concerns about Assessments in Inquiry-Based Learning Methods Courses

Inah Ko
University of Michigan

Vilma Mesa
University of Michigan

Preliminary Research Report

Abstract

We present initial findings of ongoing research that investigates the nature of instructors' concerns as they design and use assessments for their students using inquiry-based learning (IBL) approaches. Using data collected from biweekly online-teaching logs written by 39 instructors, we categorized concerns into three major themes: Item Design/Assessment, Course/Resources, and Student difficulty. We compare two areas of concerns (designing assessment and using quizzes, tests, and exams) according to the type of concern and the instructor's experience with IBL, course level, and year by using the frequencies of each category cited for each log. Our work will contribute to IBL research by analyzing instructors' challenges as a preliminary study to enhancing IBL teaching and learning in college mathematics education.

Keywords: Inquiry-based learning, Assessment design, Assessments

As the growing number of studies present the positive outcomes of Inquiry-Based Learning methods (IBL) in college students' mathematics learning (Smith, 2005; Rasmussen & Kwon, 2007; Hassi, 2009; Lauresen et al., 2011), a wide body of educators and researchers are turning their attention to IBL. The key principle of the IBL methods is based on the Moore method named after R.L. Moore, which aims for student-centered instruction in which students are encouraged to create knowledge by themselves (Coppin, Mahavier, May, & Parker, 2009).

Current interest in the inquiry-based learning methods (IBL) has sparked several investigations on how faculty navigates this kind of instruction. In particular, Laursen and colleagues report that instructors also benefited from their IBL teaching as IBL teaching enabled instructors to have deeper understanding of students and learning (Laursen et al., 2011). However, most positive results are from experimental environments where IBL instruction is designed by researchers. In addition, in adopting an innovative approach to instruction, the most problematic area that challenges even the most experienced teachers is assessment (Macdonald, 2005). Instructors who use IBL methods should be able to assess students based on a variety of student products, including written exams, presentations, and activities. This feature creates many challenges for instructors, so we focused on instructors' concerns when they design and use IBL assessments. Thus, in this study we want to find answers to the following questions:

1. What concerns or challenges do IBL faculty face when they assess their students?
2. What are the differences between the concerns expressed by faculty who indicate having little experience with the method versus faculty who indicate having more experience with the method?
3. What are the differences between the concerns expressed by faculty teaching lower division courses, upper division courses, or courses for future teachers?

Organizing and documenting instructors' concerns contributes to finding solutions to address concerns that instructors have faced with IBL and providing suggestions for practice. We believe that our work to be the first step for teachers' professional development that will lead toward effective IBL teaching and learning.

Methods

Primary source of data used in this report are entries in online teaching logs collected over two-year period from 54 instructors teaching with IBL in college mathematics courses. In the logs instructors were asked to write the challenges that they faced and the concerns they had when they design and use IBL assessments for their student as well as solutions used or planned to resolve the challenges. They submitted logs every other week for the duration of the course. Prior to recruitment, each instructor selected a level experience with IBL teaching from four possible categories (Beginner, Novice, Advanced, Expert) according to their own measure. The instructors who classified themselves as 'Beginner or Novice' are instructors who believe they have little experience with IBL. The instructors who identified as 'Advanced or Expert' are those who have had more experience with the method. We also collected course information in order to identify whether the course was intended to lower or upper division students, or whether it was intended for future teachers. Table 1 shows distribution of faculty submitting logs.

Table 1: Distribution of faculty in sample submitting logs.

	Novice/Beginner	Advanced/Expert	Total
Future Teachers	12	7	19
Lower Division	15	7	22
Upper Division	8	5	13
Total	35	19	54

To derive and develop a main theme of instructors' concerns according to their logs, we looked into each log and tagged codes to every incident that represented a concern contained in a log. We used the qualitative method of comparing incident with incident (Corbin & Strauss, 2008) in the process of coding. As we moved along with coding logs, each incident in each log is compared to previous codes for differences or similarities. If the incident is found to have similar theme to the previous coded one, we used the existing same code or grouped them together under higher dimensional codes. Otherwise, we created new codes and put the incident under a different category. After tagging all incidents with codes, we listed all codes that we created and merged some of them with more broad conceptual codes for multi-dimensional analysis.

Findings

Over the two-year period 30 instructors submitted 69 incidents in their logs on concerns they had when they designed IBL assessments and 33 instructors submitted 107 incidents on concerns related to quizzes, tests, or exams. Each incident was assigned to one of three categories of concerns: Item Design/Assessment, Course/Resources, and Student difficulty.

Designing Assessments

The categories of concerns about designing assessments and its frequencies are shown in Table 2. First, an analysis of concerns in the area of 'Designing assessments' shows that the most frequently cited category of concern was 'Item Design/Assessment' (62%, 43/69). The concern labeled 'Item Design/Assessment' refers to the challenge of creating/designing specific problems/tasks, or it can refer to concerns about the individual assessment itself in an IBL environment. Specifically, 42% (18/43) of the 'Item Design/Assessment' incidents were about designing good problems for a specific math subject or problems that accurately reflect the teacher's purpose (e.g., "Wanted to create an assignment that have more conceptual values than being computational." Log4.3_ A). Another 28 % (12/43) of these incidents were about difficulties with adjusting the level of assessment (e.g., "I am still struggling to design appropriately challenging problems that are better targeting group work." Log1.3_A). The second most frequently reported concerns in the area of 'Designing assessments' comprised concerns about 'Course/Resources.' This category represents the challenges that arise from the ways that IBL courses differ from other courses or that arise from a lack of IBL resources (23%, 16/69, e.g., supporting materials for designing the assessments, grading tools, adequate time for IBL tasks). This category encompasses a wide range of concerns caused by institutional responsibilities or course-level challenges rather than by student assessment itself. For example, 31% (5/16) of all of the concerns that were tagged in 'Course/Resources' are about grading IBL assessment (e.g., "I am having my student give oral presentations, and I wanted to determine an effective way to score these presentations. I am trying to find a way to give a fair grade based on the students work." Log2.5_B). These concerns reflect the need of a fair scoring system that integrates various types of IBL assessments, including oral exams, student presentation, and take-home exam. We assigned these concerns into the category of 'Course/Resources,' because the solution to these would require course-level improvements and resources rather single assessment-level solutions. In addition, 25% (4/16) of all comments were about time management, that is, the teacher's challenges in managing time for IBL-based teaching.

Finally, 14% of the comments (10/69) about 'Designing assessments' concerns among all of the tagged incidents in the logs over two year period were about student difficulty/level. We defined this category as the concerns that instructors see their students are experiencing as they take an IBL assessment or understand the concept of IBL.

Table 2. Frequency of Concerns on Designing Assessments (N = 69)

Category	Frequency (%)
Item Design/Assessment (Desigining good problems, Appropriate level)	43 (62%)
Course/Resources (Grading, Time management, lack of resources)	16 (23%)
Student Difficulty/Level	10 (14%)
Total	69 (100%)

Using Quizzes, Tests, or Exams

In contrast with the concerns about 'Designing assessments,' the most frequently cited category of concern in the area of 'Quizzes, tests, or exams' among the three major categories comprises concerns about 'Student Difficulty/Level.' Fifty-six percent (60/107) of the incidents were about the students' poor performance on quizzes and exams or concerns about the students' lack of preparation (e.g., "Students just hadn't prepared well enough for the exam," Log3.7_D). Thirty percent (32/107) of the tagged concerns were about 'Item Design/Assessment', that is, the challenges that teachers encounter while creating various types of IBL-based assessments (e.g., balancing question types, creating multiple-choice questions or take-home exam). The remaining 14% (15/107) were concerns about 'Course/ Resources'. A typical example of a course-related concern is the increase in the teacher's workload (e.g., "These quizzes and exam corrections have been time-consuming, and would not be possible without the help of an excellent TA this semester." Log2.3_E). See Table 3.

Table 3. Frequency of Concerns on using Quizzes, Tests, or Exams.

Category	Frequency (%)
Student Difficulty/Level	60 (56%)
Item Design/Assessment	32 (30%)
Course/Resources	15 (14%)
Total	107 (100%)

Our analysis suggests two trends. First, Beginner or Novice (BegNov) instructors have more concerns about IBL-based assessments than do the Advanced or Expert (AdEx) instructors, and AdEx instructors have more concerns about their students (student performance, difficulty) than about themselves as compared to the BegNov instructors. Second, instructors who are teaching low division (LD) courses tend to have more concerns in general than do the instructors who are teaching courses for future teachers (FT) or upper division (UD) courses. As a next step in our analysis, we will be determining whether these differences between groups are statistically significant.

Discussion

Overall, regarding concerns about 'Quizzes, tests, or exams', most instructors were worried about the students rather than about themselves or external resources. For example, instructors were more concerned about their students' preparedness or about the students' poor abilities in the new IBL environment than they were about their own preparedness or their ability to lead IBL instruction. However, relatively few instructors mentioned the students' low performance or difficulties as concerns that they have when they design IBL assessment. During the design stage, they focused more on their knowledge of creating/designing good problems. This result might be consequential to the nature of the questions that we used. In other words, the pre-determined name of area 'Designing assessments' might cause instructors to focus on themselves, because the main agent of designing is the instructors, whereas 'Quizzes, tests, or exams' reminds them of more of the general difficulties in solving IBL problems. In addition, many instructors seem to have concerns and feel pressure due to their view of what is 'IBL'. Most of them feel very much inclined to use 'IBL' assessments (although this definition is not very clear), but regard these assessments as conflicting with standard assessments (define) or requiring too much additional work.

Questions for the Audience

1. We categorized instructors' concerns by merging similar concerns and creating higher level of code. What could be other conceptual layered codes that could be used to capture essentials from logs as well as differences between concerns?
2. We collected not only concerns but also solutions. However, most solutions instructors wrote were too generic (e.g., "spend too much time" or "find a good resources") or not useful ("I have no idea"). What types of questions could be added to the logs so that we can utilize collected solution data for further analysis of their concerns or their perception about IBL assessments?

References

Corbin, Juliet, and Anselm Strauss. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 3rd ed. Los Angeles, Calif.: SAGE Publications, Inc, 2008.

Coppin, C. A., Mahavier, W. T., May, E. L., & Parker, G. E. (2009). *The Moore Method: A path way to learner-centered instruction*. Washington, DC: The Mathematical Association of America.

Hassi, M.-L. (2009). *Empowering undergraduate students through mathematical thinking and learning*. Paper presented at the 15th International Conference of Adults Learning Mathematics, Lancaster, PA.

Laursen, S., Hassi, M. M. L., Kogan, M., Hunter, A. B., & Weston, T. (2011). *Evaluation of the IBL mathematics project: student and instructor outcomes of inquiry-based learning in college mathematics*. Colorado University.

Macdonald, R. (2005). 9. Assessment strategies for enquiry and problem-based learning. *Handbook of Enquiry & Problem Based Learning*. Barrett, T., Mac Labhrainn, I., Fallon, H. (Eds). Galway: CELT, 2005.

Nardi, E., Jaworksi, B., & Hegedus, S. (2005). A spectrum of pedagogical awareness for undergraduate mathematics: From 'tricks' to 'techniques'. *Journal for Research in Mathematics Education*, 36(4), 384-316.

Rasmussen, C., & Kwon, O. (2007). An inquiry oriented approach to undergraduate mathematics *Journal of Mathematical Behavior, 26*, 189-194.

Smith, J. C. (2006). A sense making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior*, 25, 73-90.

Deploying Tasks Assessing Mathematical Knowledge for Teaching as Tasks for Professional Preparation

Yvonne Lai
University of Nebraska-Lincoln

Heather Howell
Educational Testing Service

Abstract. Mathematical knowledge for teaching (MKT) has been shown to be a measurable construct impacting instructional quality and student outcomes. The primary examples that educators have for MKT tasks are those that were designed and validated for assessment purposes. It is not known to what extent features of a task that support its use as assessment may support or hinder its use in instruction. We examine this tension by studying the use of two such MKT tasks in a course for prospective teachers. Key considerations for using MKT tasks in professional preparation tasks were how the MKT task represents teaching practice and the possible purposes of using that representation in teacher education.

Purpose

In recent years, evidence has been amassing that K-12 mathematics instruction is impacted by teachers' knowledge (e.g., Baumert et al., 2010; Hill, Rowan, & Ball, 2005; Rockoff, Jacob, Kane, & Staiger, 2011). These studies, demonstrating positive relationships between teachers' knowledge and student outcomes, point to the importance of *mathematical knowledge for teaching* (MKT) (Ball, Thames, & Phelps, 2008) as a form of professional knowledge of content. However, the relationship between MKT as a measurable construct and MKT as a learnable, teachable body of knowledge requires more study if teacher education programs are to prepare teachers in MKT (National Academy of Sciences, 2010).

Assessment of knowledge relates to – but is distinct from – knowledge development. One reason why it may be difficult to bridge the gap between research on MKT and the practice of teacher education is that prevalent examples of MKT, such as the Learning Mathematics for Teaching instrument (e.g., Learning Mathematics for Teaching, 2008) were originally designed for assessment, and were often generated based on data from practicing teachers. This situation suggests potential difficulties for the use of such MKT tasks in preparation programs; the tasks may not match with prospective teachers' perceptions of teaching practice, and features of the tasks that are valuable in assessment may be less appropriate for teacher education. Our research addresses the questions: *What features of an MKT assessment task, which represent teaching practice, can be support for using the task in instruction? What about these features can be in tension with using the task in instruction?*

Theoretical and conceptual perspectives

This paper takes the perspective of the theory of mathematical knowledge for teaching outlined in Ball, Thames, and Phelps (2008), which continues earlier work by Shulman (1986) building an understanding of the professional knowledge base required by teaching. *Mathematical knowledge for teaching* is the mathematical knowledge required to carry out the recurrent work of teaching. Such work may include, for example, analyzing student work, giving an explanation, or selecting an example.

Tasks created for MKT assessment often include pedagogical context: the task presents a teaching scenario whose features are intended for use in solving the task (e.g., Baumert et al., 2010; Hill, Schilling, & Ball, 2004). This context constitutes much more than just window dressing for a traditional mathematics problem, and in fact situates the test taker in ways that partially define the construct being tested. Features of this context include the content, students, and instructional purpose; such features orient the test taker to approach the task as a legitimate problem of teaching (Lai, Jacobson, & Thames, 2013; Phelps, Howell, & Kirui, 2013).

To know that an MKT task functions well as part of an MKT assessment, it must be validated as capturing professional knowledge for mathematics teaching. The MKT tasks discussed in this paper were designed for assessment under the Bill and Melinda Gates funded Measures of Effective Teaching (MET) project (Bill and Melinda Gates Foundation, 2012), and validated under Kane's (2006) validation framework (Gitomer, Phelps, Weren, Howell, & Croft, in press). Kane (2006) specifies, however, that both target population and intended use are key components of a validity argument, suggesting it would be problematic to assume tasks validated for assessment uses with in-service teachers are valid for instructional uses of either in-service or preservice teachers. However, these and other similar assessment tasks (for example, the Learning Mathematics for Teaching tasks) currently constitute the main body of examples used by educators to understand MKT.

We investigate the use of MKT tasks, whose answer requires bringing MKT to bear on the solution process, as a resource for teaching and learning (c.f. Brousseau, 1997). How a task is set up by an instructor to use prospective teachers' prior knowledge and experiences, as well as how the task is enacted, that is, how an instructor directs and scaffolds the prospective teachers' work, impacts what prospective teachers learn from working on an MKT task (Stein, Grover, & Henningsen, 1996). The set up and enactment consist of instructional interactions: the interactions between the content, the instructor, the learners, and the environment. This study uses the conception of teaching as the management of instructional interactions between teacher, content, students, and the environment (Cohen, Raudenbush, & Ball, 2002). We use *intended MKT task* to refer to the original version of a task, and *enacted MKT task* to refer to the version of the task used in instruction (cf. Stein et al., 1996).

Data and Method

To examine the affordances and tensions that MKT tasks designed for assessment can bring to instruction, particularly the management of interactions between instructor's purposes, the MKT tasks, and teachers, we used data from a semester-long mathematics methods course for prospective high-school teachers which emphasized MKT and during which MKT assessment tasks were used by the instructor as curricular material. Primary sources include audio recordings of the lessons, planning notes, and slides used in the 14 lessons that semester. Secondary sources include pre and post lesson interviews with the instructor who taught the course. The instructor used MKT assessment tasks from the MET project in 2 lessons, and in 9 others used MKT tasks of her own design based on her exposure to MET and LMT generated tasks. For clarity, in this paper we will refer to the instructor who taught this course, who is one of the authors, as *instructor*, to the prospective teachers enrolled in the course as *prospective teachers*, to the named teacher described in an MKT task as *teacher* and to the implied students described by an MKT task as *students*.

Design of Analysis: Interaction between Instructional Purpose and MKT Task

In this paper we discuss two cases of the interaction between an MKT task and the instructor's purposes, illustrating how the interactions of these played out in instruction in ways that reveal both the tensions and the affordances around the use of such materials. The two cases are oriented around the MKT tasks titled "Kane" and "Anderson", named for the hypothetical teacher in the task scenario.

As Hiebert and Grouws (2007) write, teaching consists of "classroom interactions among teachers and students around content directed toward facilitating students' achievement of learning goals'' (p. 372). How learning goals are emphasized and tasks are managed influence opportunities to learn (Hilbert & Grouws, 2007). Sleep (2012) introduced the notion of "steering instruction towards the mathematical point" – work a teacher does to ensure that students are engaged in intended mathematical work that serves the learning goals. We use the notion of "steering instruction" toward intended work with MKT. Interaction between teacher purpose and curriculum material plays out in instruction – our point is to illustrate specific ways in which these interact when the curriculum material is MKT problems and to analyze whether there are specific attributes of either problem or purpose that make certain types of interactions more or less likely.

We consider the MKT task itself to be one input to each episode of instruction; a set of codes was developed to describe each MKT task in terms of its implicit purposes, types of student work represented, and other relevant features. All MKT tasks utilized in the course were coded, but our analysis is restricted to the two lessons in which MKT tasks originally drafted for assessment purposes were used in their original form. This restriction is partly motivated by our observations that these tasks, because they were drafted for a specific and different purpose, have more clearly defined descriptors independent of their use in this course. Other MKT tasks, because they were adapted or created by the instructor to begin with, are more difficult to analyze in terms of how their original form interacted with the instructor's purpose because they were not utilized in an original form that was entirely free of adjustment. It is also partly motivated by practical concerns. Assessment tasks are the form in which much recent work around MKT exists and is publicly available, and an understanding of how these existing resources can be of use in teacher education opens the door for a rich set of resources to become available for use in this way.

We consider the instructor's purposes to be a second input to each instructional situation. The instructor's purposes are described in detail, based on the data records, and were similarly coded descriptively as to the nature of each purpose. Finally, we look at the outcomes, which are the episodes of instruction as they played out in the semester course. We coded transcripts of the course for evidence of interactions between the instructor's purposes and the underlying features of the MKT tasks. We take it as given that there are interactions between the MKT tasks and the instructor's purposes – our goal here is not to provide evidence *that* the teacher's purpose interacts with curricular resources – rather it is to describe the nature of these interactions, with particular attention to points of tension and to interactions that are suggestive about the usefulness of MKT tasks more generally.

Example MKT Task and Episode of Instruction

We find it helpful in describing our methods to ground the discussion in an example from the data. In the section that follows, we discuss the Kane task, an MKT task used in the course, and how instruction around this task unfolded. We then use this example throughout to clarify our coding and analytical decisions.

The Kane task is typical of other MKT assessment tasks in its use of an instructional scenario to set the stage and in its presentation of records of practice, in this case examples of student work, as stimulus for the test taker to respond to. The work to be done here is mathematical, as the task is to determine whether the reasoning is correct; this mathematical work is situated in a teaching context. Teachers and non-teachers alike might be called on to solve equations correctly. Only teachers, however, would be called on to evaluate unconventional solutions techniques produced by others, particularly in a context where the presentation of the work might be unclear or incomplete. We begin by describing the task itself, then present a summary of the data collected in the form of a teaching episode in which the instructor used this task in a course for prospective teachers.

The Kane task. The Kane MKT task (see Figure 1) asks the reader to analyze five given examples of student work, all of which lead to a correct solution. For each example, the essential question to answer is whether the steps shown, which vary in detail and clarity, correctly lead to the given answer.

The student work shown in options (A), (C), and (E) is correct. The approach shown in (A) typifies a conventional approach, although the variable is on the right rather than left hand side of the equation just prior to solving for x. However, such a choice is often made by students in order to make it possible to work with positive rather than negative coefficients when solving for a variable. Not all steps are shown; for example, the test taker has to infer that 5x has been added to each side of the equation in the first step. However, valid arithmetic properties can be used to justify each line of the student's work. The work in (C) is more detailed. The student has subtracted 13x and added 10 to both sides of the equation, and prior to solving, the variable terms are expressed in the left hand side of the equation. The student completes solving the equation by multiplying by –1/18, which is correct although probably not the most common student approach. The student in (E) solves similarly, showing work and using routine notational convention to indicate additive inverses.

Option (B) is more difficult to follow as less work has been shown, but it appears the student has added "unlike" terms on each side, for example, obtaining the sum $3x$ from adding $-5x$ and 8. Although the student coincidentally arrives at the correct answer, the solution is not correct. Option (D) shows a similar mistake. In this case, the student combined unlike terms only on the right hand side, again arriving at a correct answer only coincidentally.

During a lesson on solving multistep equations, Ms. Kane asked her students to solve the equation $-5x + 8 = 13x - 10$. While walking around the classroom looking at what the students were writing, she noticed several different strategies. For each of the following student solutions, indicate whether or not the work provides evidence that the student is reasoning correctly about this problem.

(A)

$$-5x + 8 = 13x - 10$$
$$8 = 18x - 10$$
$$18 = 18x$$
$$1 = x$$

(B)

$$-5x + 8 = 13x - 10$$
$$\frac{3x}{3} = \frac{3}{3}$$
$$x = 1$$

(C)

$$-5x + 8 - 13x + 10 = 13x - 10 - 13x + 10$$
$$-5x - 13x + 8 + 10 = 0$$
$$-18x + 18 - 18 = 0 - 18$$
$$\frac{\cancel{1}}{\cancel{-18}} \cdot \frac{\cancel{-18}x}{\cancel{1}} = \frac{\cancel{-18}}{\cancel{1}} \cdot \frac{\cancel{1}}{\cancel{-18}}$$
$$x = 1$$

(D)

$$-5x + 8 = 13x - 10$$
$$-5\cancel{x} + 8 = 3x$$
$$+\cancel{8}x \qquad +5x$$
$$\frac{8}{8} = \frac{8x}{8}$$
$$1 = x$$

(E)

$$-5x + \cancel{8} = 13\cancel{x} - 10$$
$$-13x\ \cancel{+8} \qquad -\cancel{13}x\ -8$$
$$\frac{-18x}{-18} = \frac{-18}{-18}$$
$$x = 1$$

Figure 1. The Kane Task

4

The common student mistake showcased in this MKT task is that of "combining unlike terms". This misconception is often related to purely procedural instruction in which students learn as a rule that addition can be performed only between "like" terms but that multiplication can be performed on “unlike” terms. The underlying mathematical idea is that combining "like" additive terms is possible because terms are "like" when their variable factor is the same, and this variable factor (here, x), can be factored (making way for the use of the addition operation on coefficients) and then distributed. The records of practice (student work) presented also provide a space for thinking about a number of issues that an instructor might raise if sharing this MKT task with a class of prospective teachers. An instructor might use the opportunity to focus on diagnosing not just the errors but also the non-errors; what do the students who made mistakes likely already understand? The three correct solutions vary from one another in mathematical structure, providing an opportunity to discuss what makes a solution strategy conventional or unconventional, whether or in which contexts this should matter, and whether solutions strategies are fully general or apply to only subsets of problems. They also provide examples of approaches students might typically take, and illustrate that even simple problems can elicit correct student responses sufficiently complex that a teacher must work to make sense of them. The variation in how much work is shown is also typical of students, and provides an opportunity to think about correctness versus completeness of solutions, the difficulties teachers encounter in trying to follow student work, and the use of notational conventions (such as, for example, writing the division of both sides on the equation in a single step) that help or hinder the reader of a solution in trying to follow the work. Activity with this MKT task could also allow for discussion of how one can diagnose, on the basis of written work, what a student might be doing more generally.

In addition, the embedded instructional purpose implies certain values about teaching. For example, implicit in the task situation is the idea that teachers should be examining student solutions at the individual level, and that understanding what students might understand or not understand is important work of teaching. Implicit in the scoring (that the work in (A), (C), and (E) "provide evidence that the student is reasoning correctly", and that the work in (B) and (D) does not) is that a solution that is correct but not detailed can be sufficient, at least in the instructional situation outlined in the task.

That artifacts of student work are presented in the task has potential to be both useful and problematic in using such a task with novice teachers. Examining student solutions to a mathematics problem is unquestionably part of the routine work of teaching mathematics, and the authenticity of this particular task situation hinges on how authentic the artifacts of student work are. If seen as authentic, the artifacts are examples of the types of things that students do with such problems and may help novice teachers to appreciate the range of responses and what it takes to respond to them. On the other hand, because the teachers may have little experience with student work, the situation may not ring true if they do not believe that students actually take such approaches. This concern is supported by analysis of MKT assessments showing that MKT tasks associated with analyzing student work differentiate more strongly between in-service and pre-service teachers, suggesting that pre-service teachers may not have the necessary MKT to address such tasks, or to be able to approach the problems from a teaching perspective (Phelps, Howell, & Schilling, 2013).

Vignette: The Kane task in instruction. What happened when the instructor used this task in a course for prospective teachers?

At the beginning of the lesson, the prospective teachers worked on warm-up problems, one of which was the Kane task. The prospective teachers were asked first to respond to the student-level task “solve the equation $-5x + 8 = 13x - 10$” and then to respond to the Kane task itself. The instructor had planned for the Kane task to serve as a context for practicing the work of assessing and diagnosing student understanding. Additionally, she wanted teachers to learn to frame diagnoses of student understanding in terms of underlying mathematical properties and laws. Her reasoning for this was that if the prospective teachers could learn to view routine exercises such as solving $-5x + 8 = 13x - 10$ as involving procedures that have a conceptual basis, perhaps they would be able to help their future students do the same, thus heading off some of the errors shown in the task.

As the prospective teachers' discussion wound down, the instructor called the class to attention. Before launching into the mathematical diagnoses, she wanted to make sure that the prospective teachers would work from a common understanding of which solutions were correct and which were incorrect. The instructor asked the prospective teachers, "Are there any questions about the fact that (B) and (D) show incorrect reasoning, and (A), (C), (E) show evidence of correct reasoning?"

Unwilling to judge the students' work as either "correct" or "incorrect," Marisa said that the student showed "*some* understanding." Karen agreed: "It looks like they knew to move, to add, the terms to one side. They got rid of the –10 and then got $3x$. I thought maybe it was a copy error so they *did* understand but they just had the wrong problem. So it was correct reasoning, but the problem was incorrect. That's how I saw it." In considering what a student might have understood, Marisa and Karen concluded that there may have been at least partial understanding. The options available to choose from -- "does provide evidence" and "does not provide evidence" – were not sufficiently nuanced for the point they wanted to make, which was that there is *some* evidence that *some* reasoning is correct and *insufficient* evidence to conclude that the reasoning is definitely incorrect.

However, Marshall responded, "It could well have been a copy error, like they rewrote the line incorrectly, but I think what's more likely is that they had a misconception about like terms. What I think is likely is that the students combined $13x$ and –10 and got $3x$ out of that. So I think *that* shows there's not complete evidence that the student shows understanding." Marshall later noted that this misconception "happened on larger scale" with $5x$ and 8.

Jonathan commented that perhaps all the students were "getting to the right answer" because the solution to the equation is $x = 1$.

At this point the instructor transitioned the discussion from answering the Kane task to an extended purpose of using Ms. Kane's students' work to diagnose the underlying mathematical laws or properties that the students might have misunderstood. She had hoped that the prospective teachers would see that incorrectly adding "unlike" terms is related to students not thinking about summing "like" terms as an application of the distributive property, and that this might serve as an example of how attention to the underlying laws can actually provide a conceptual basis for understanding simple algebraic procedures that colloquial but succinct descriptions like “combine like terms” lack.

Several prospective teachers pointed to procedural aspects of the students' work. Melissa commented about the student work in (D) that "they know how to add the same thing to both sides, so they add $5x$ to both sides, and then they divide both sides by 8." Isabelle added that, "Along with that, I think the student knows to get like terms on both sides."

One prospective teacher pointed at a difference between mathematical properties of addition and multiplication. "I think that students might be confused because you can multiply $3x$ times 2, but you can't add $3x$ plus 2. So they can get confused – 'Why can't I add and subtract like terms?'" This comment begins to touch on the idea that the instructor was aiming for, but still in colloquial language more like that a student would use, and not in the mathematically precise language the instructor was hoping for.

The instructor reframed the prospective teachers’ comments in terms of mathematical principles, saying that the students perhaps understood inverse operations, but not necessarily how to add the terms in the equation. She then pointed at the prospective teachers' activity throughout their work on the Kane task, saying, "How we did this was put ourselves in the students' shoes. We thought about what they've seen before, what steps they took, and why. Another thing we did was look for patterns. Finally, we thought about common rules and how they might have not been applied correctly. ... Three common rules that are often misapplied are distributive property, adding fractions, and simplifying expressions. We just saw these." The instructor then moved on to another activity.

In what ways did the instructor’s purposes interact with the MKT task in the enactment of instruction? Clearly, the instructor had multiple purposes in using the task. One was simply to engage the prospective teachers in answering the task as written, a task that requires analyzing the student work samples to decide if they are correct. Another related purpose was to engage them in practicing the skill of diagnosing student thinking, including what students *do* understand as well what they *don’t*. And a third purpose was to engage the prospective teachers in thinking about the underlying mathematical properties that justify the procedures the students were misapplying. All three purposes make sense in the context of the given MKT task, although only the first matches the purpose of the task as originally presented as an assessment task. Why is the second purpose achieved by the instructor but the third is not? While a number of factors are certainly at play in the instructional episode, our focus is on the interaction of the instructor’s purposes and the task features. Did the framing of the student work, for example predispose the prospective teachers to talk about the work samples in student-friendly language, undermining the instructor’s efforts to focus them on mathematical properties expressed in mathematical language? We examine in this paper, more generally, what features of an MKT assessment task support or limit the task's use in instruction.

Coding Pedagogical Context of MKT Tasks

We define the pedagogical context of an MKT task as the elements of teaching and learning provided in the text of the task (Lai et al., 2013; Phelps, Howell, and Kirui, 2013). Following Lai, Jacobson, and Thames and Phelps, Howell, and Kirui, we analyzed the text of each MKT task to identify the purpose of the actor portrayed in the task (usually a named teacher), instructional records of practice such as student work, student background, and the way in which instruction as portrayed in the MKT tasks is organized; these elements, as a set, capture the pedagogical context in tasks including those of LMT and MET assessments (Lai, Jacobson, & Thames, 2013). Table 1 provides a summary of the codes used in this analysis, using the Kane task as an example when possible.

Table 1. Codes used in this analysis.

Pedagogical context code	Description	Example
Purpose of the actor portrayed in the task (usually a named teacher)	Purpose within MKT tasks for actor or hypothetical teacher in the pedagogical context of the task, as stated or implicit in the task text.	Determining whether or not the given samples of student work provide evidence that each student is reasoning correctly about the task.
Primary records of practice include records to which other (supplementary) records respond. Cases where only one record is presented are coded as primary.	Primary and supplementary records of practice are distinguished by whether one depends on the other. For example, if the scenario describes a problem the teacher assigned and then shows student responses to the problem, the problem is coded as a primary record and the student work as supplementary records.	Asked her students to solve the equation: $-5x + 8 = 13x - 10$.
Supplementary instructional records of practice	Records in response to a primary instructional record, e.g., student work on an assigned task where the task was also given.	Student work on the task, e.g., $-5x + 8 = 13x - 10$ $8 = 18x - 10$ $18 = 18x$ $1 = x$.
Student background	Background information given, possibly including information about when in a unit of instruction the given instructional moment occurred, the grade level of the students, or particular concerns the teacher might have.	No example provided in the Kane task. The instructor later uses the Kane task students' misconceptions as a premise for an activity. For this activity, the students misunderstanding of aspects of solving multi step equations is information about the student background.
Organization of instruction	Information about the nature of instruction, e.g., involving group work, circulating, discussion.	Walking around the classroom looking at what the students were writing.

Coding Instructional Purposes of MKT Tasks

We define an instructional purpose as a purpose – such as a message that prospective teachers were to take away or the reason for doing something – stated by the instructor about an MKT task, either in the text of planning notes or spoken aloud to the prospective teachers in instruction.

One of the instructor's purposes associated with the use of the Kane task, for example, mirrors Ms Kane's purpose as represented in the pedagogical context: "to practice determining whether students' work provides evidence of correct reasoning." The instructor states this as a goal for the class in her planning notes and interview about this lesson. A closely related purpose was for prospective teachers "to practice analyzing student thinking by using sample student work to consider what the student might understand and not understand." This goal statement was extracted from the instructor's statement to the prospective teachers, "Now we'll get to student thinking. So the goals here are to practice analyzing, assessing, responding to, and soliciting student thinking. We're going to start with analyzing student thinking by using a sample of student work." The planning notes further specified that the "goal" of the Kane task was to "use a sample of student work to consider what the student might understand and not understand."

Another purpose associated with the use of the Kane task was for prospective teachers "to understand that common mis-applied 'rules' include distributive property, especially with negatives and variables; adding fractions, especially with variables; and simplifying expressions of the form $ax+b$ or $a+bx$." When the instructor concluded the activity on the Kane task, she stated, "We thought about common rules and how they might have been applied correctly ... Here are three examples... the distributive property. Adding fractions. Simplifying expressions. This is what we just saw." The accompanying slide gave more detail about each of these "examples" of commonly mis-applied rules.

After coding the data for evidence of the instructor's purposes for use of each of the MKT tasks, we then described these purposes with brief text descriptors and analyzed the set of descriptors for trends. We noticed five distinct kinds of purposes for the instructor's use of MKT tasks. Below we describe each type. Table 2 summaries these five kinds of purposes with examples.

Knowledge of mathematics, mathematics teaching, or pedagogy. This code describes cases where the instructor's purpose in using the MKT task is specific to the content (mathematical or pedagogical) that it showcases. For example, the task might emphasize attributes of mathematical topics or illustrate the difference between procedural and conceptual approaches.

Modeling good teaching practice. This code is applied to cases in which the instructor uses the MKT task as an opportunity to model good teaching practice, either by showcasing good teaching by the implied teacher of the task or by herself modeling practice around the task. For example, the task context might model attentiveness to unconventional student solutions on the part of the teacher even if it is not explicitly stated that the implied teacher's goal is to be attentive to unconventional solutions.

Norms and responsibilities to advocate. The instructor uses the MKT task as part of an argument, generally explicit, for particular norms or responsibilities of the teaching profession.

Practice by prospective teachers. This code describes cases where the instructor's intention is that the students practice some skill, whether mathematical, pedagogical, or both.

Showing that a situation can or does occur in teaching practice. This code describes instructor purposes that indicate the instructor is using the MKT task or some component of it as evidence that something can or does happen. This includes purposes where what is being illustrated is types of student responses to a situation.

Table 2. Types of instructor purposes associated with instruction using MKT tasks.

Purpose type	Examples
Knowledge	That "gathering like terms" is an application of the distributive property.
Model	To model how a teacher might think about a problem from the student's point of view in order to understand the student's work.

Norms and responsibilities	To argue that focused questions are important because they are a resource for helping students build confidence and see their own progress.
Practice	To practice analyzing student thinking by using sample student work to consider what the student might understand and not understand.
Show	That students misapply "rules" such as the distributive property, especially with negatives and variables; adding fractions, especially with variables; and simplifying expressions of the form $ax + b$ or $a + bx$.

Coding Role of MKT Task in Instruction

To investigate the role of MKT tasks in instruction, we analyzed the use of the MKT task by the instructor. In particular, we examined instructional interactions between the instructor's purposes, the prospective teachers' actions, and the MKT task in which the use of the pedagogical context of MKT tasks was visible. Such interactions were coded according to the ways in which the task was an *affordance* relative to the instructor's purposes, and for the ways in which there was *tension* visible in the instructor's use of the task. In many cases, an interaction was considered to have both affordances and tensions. Our interest was less in whether an interaction was coded for affordance, tension, or both, and more in characterizing the ways in which these affordances and tensions played out and what features of the MKT task and of the instructor's purposes might influence the interaction.

Example. An example of an interaction can be seen in the class's work with the Kane task. The records of practice (student work samples) from the Kane task engaged prospective teachers; prospective teachers evaluated the hypothetical students' work as intended and cited specifics in supporting their claims about what the students did or did not understand. The prospective teachers even went further than required by the task as written by using the records to determine the ways in which students demonstrated understanding or not, and used them as a common reference from which to hear and respond to other prospective teachers' contributions. The instructor was later able to use this discussion about the records of practice to introduce a heuristic for diagnosing student work that included considering what a student may or may not understand. This supported one of the instructor's purposes, that the prospective teachers *practice* diagnosing student work. On the other hand, the instructor's *knowledge* purpose was for the prospective teachers to know the specific underlying mathematical laws. Although the prospective teachers identify where evidence of lack of understanding occurs in the records provided by the MKT task, their descriptions of these locations tended to be procedural (e.g., "They know to get the x on one side", "They know to add the same thing to both sides") rather than statements about exactly which mathematical laws or properties are being used in valid or invalid ways. Each procedural error could be described in terms of the mathematical laws or properties being misused but the instructor does not seem to be able to move the prospective teachers' contributions toward the end of identifying specific mathematical laws. This example illustrates how interactions between the instructor's purposes and the MKT task can simultaneously be cases of *affordance* and *tension*, often because there are multiple purposes layered on a single activity as enacted in instruction. What is of interest here is the ways in which these are *affordances* and *tensions*, which is explained in the following section.

Instructional use of pedagogical context as affordance for an instructor's purpose. We analyzed the transcripts, audio records, and visual records to determine whether and how instruction used elements of pedagogical context towards a purpose of the instructor. Recall that "instruction" refers to interactions between the instructor, the prospective teachers, and the MKT task. Here, by "used toward a purpose of the instructor", we mean that a counterfactual absence of the element would remove the context for statements made by or actions taken by the instructor or prospective teachers that supported the purpose of the instructor, i.e., what happened would not have been likely to happen absent the specific contextual element. Such actions might include those that generated conviction towards the purpose or promoted engagement with activities that supported the purpose. In the above example, the records of student work were clearly necessary in order for the prospective teachers to engage in the instructor's purpose of them practicing diagnosing student work. We also argue that an implicit purpose of the Kane task, Ms. Kane's desire to determine whether students understand, is an element used in instruction, as the situation would make little sense without this contextual framing.

We also recorded whether a modified version of the element was used in this sense; in these cases, we also recorded a description of the modified element. In coding *whether* an element of the pedagogical context of an MKT task was used for a purpose, we used Y, XI, XT, N, and O to denote categories of use. Table 3 summarizes these codes. Additionally, we described *how* the instructor and prospective teachers' interactions with elements moved instruction towards the purpose. The records of practice in the example above moved instruction toward the instructor's purpose by providing material that engaged the prospective teachers and offered them a reference point for discussion. The implicit purpose may have moved instruction toward the purpose in a more subtle way, by creating a scenario in which the close examination of student work is reasonable and plausible, allowing the prospective teachers to go further and analyze the student work samples more deeply than was required to answer the MKT task itself.

Instructional use of pedagogical context as source of tension for an instructor's purpose. We characterized the instructional use of the pedagogical context as a source of tension analogously: if a counterfactual absence of the element would remove the context for statements and actions that moved away from a purpose of the instructor. We described *how* the instructor and prospective teachers' interactions with elements moved instruction away from the purpose. In reviewing our descriptions, it seemed that these situations had a common structure: the interactions seemed to position related but distinct aspects of instruction against each other. We then described the aspects of instruction positioned against each other. In the example above, the use of records of practice (student work samples) written in simulated student handwriting and presented as authentic student work supported the prospective teachers' engagement but may also have supported their use of student-like language to describe the student thinking, using phrases like "get x on one side" because this is the way the envisioned students would think about this step. But by doing so, the records may be positioned in opposition to the goal of describing the underlying mathematical laws, simply because this is not the way one might generally approach the task of dealing with authentic student work from real students.

In the next section, we discuss affordances and tensions that can arise in instruction when using MKT tasks. We then discuss the ways that the assessment-related features of MKT tasks – the elements of pedagogical context and their use in solving the intended MKT task – each can combine with instructional interactions as affordances and tensions relative to the instructor's purposes.

Table 3. Categories of use of elements of pedagogical context in instruction.

	Categories of use of elements of pedagogical context of MKT task and their meanings
Y	The instructor or prospective teachers used an element of pedagogical context as provided by the MKT task
XI	The instructor used an introduced or modified element of pedagogical context as provided by the MKT task
XT	The prospective teachers used an introduced or modified element of pedagogical context as provided by the MKT task
N	Neither the instructor nor prospective teachers used an element of pedagogical context that was provided by an MKT task
O	This element of pedagogical context was not provided by the MKT task and no such element was introduced by either the instructor or prospective teachers.

Results: Instruction and MKT Tasks in Professional Preparation

Our analysis suggested several affordances and tensions that can arise in instructional interactions around an MKT task, and that elements of pedagogical context contribute to these affordances and tensions. In this section, we present an additional vignette describing how the Anderson task played out in instruction. We then describe affordances and tensions, and finally we discuss the role of elements of the pedagogical context of MKT tasks in instruction.

Vignette: The Anderson Task in Instruction

Instruction in this lesson used the Kane and then Anderson task. Instruction with the Kane task was described previously. Here we describe instruction with the Anderson task (Figure 2).

Mr. Anderson asked his students to simplify the following algebraic expression.

$$\frac{2(a+1)}{3a} + 3 - \frac{2}{3a} - \frac{6a-2}{6} =$$

One of his students gave the incorrect solution shown to the right.

Of the following descriptions, which best characterizes what is wrong with this student's work?

(A) This student used the distributive property incorrectly.
(B) This student confounded mixed fractions with factors.
(C) This student forgot to cancel common factors in several places.
(D) This student needs to apply a more formal procedure by finding the common denominator and then adding all terms.

$$\frac{2(a+1)}{3a} + 3 - \frac{2}{3a} - \frac{6a-2}{6}$$
$$= \frac{2a+2}{3a} + 3 - \frac{2}{3a} - \frac{6a}{6} + \frac{2}{6}$$
$$= \frac{2a}{3a} + 2\frac{6}{6} + \frac{2}{6} - a$$
$$= \frac{2}{3} + 2\left(\frac{6}{6} + \frac{1}{6}\right) - a$$
$$= \frac{2}{3} + 2\frac{7}{6} - a$$
$$= \frac{4}{6} + \frac{14}{6} - a$$
$$= \frac{18}{6} - a$$
$$= \underline{\underline{3-a}}$$

Figure 2. The Anderson Task as Intended

Anderson task as intended. The Anderson task asks the person solving it to characterize an error made in a sample of student work on an algebraic expression. The task statement is shown in Figure 2.

To solve the Anderson task, one must take stock of the student work line by line, deduce the reasoning that the student might have taken, and then evaluate the way the student may have employed mathematical properties in valid or invalid ways. Like the Kane task, the student work "skips" steps, particular between the second and third line, right before the error occurs; however, the expressions in the second and third line are equivalent, so there is no evidence in these beginning lines of lack of understanding. Note that in the third line, the student has – correctly – expressed the quantity "3" as the mixed fraction "2 $^6/_6$." The fourth line contains the error, where the notation for mixed fractions has been – incorrectly – interpreted as notation for multiplication. Had this expression actually represented multiplication, then the notation 2 $^6/_6$ + $^2/_6$ would have represented a quantity equivalent to the product 2($^6/_6$ + $^1/_6$), a usage of the distributive property that would have been mathematically valid. Thus option (A) does not accurately describe the source of error. However, the expression does not represent multiplication, and therefore the option that best characterizes "what is wrong with this student's work" is option (B): "This student confounded mixed fractions with factors."

Although the student's work does show several fractions whose denominator and numerator have common factors, there are no errors that result from using the fractions as is, so option (C) does not capture the error. Finally, the error made in the work does not directly stem from the absence of finding common denominators as much as it does the notation for fractions and products, so option (D) does not capture the error either.

Anderson task as enacted. The instructor modified the Anderson task to ask prospective teachers to predict how student misconceptions might surface instead of asking to diagnose a specific instance of student error. The instructor displayed a truncation of the Anderson task on a slide, reading:

> Mr. Anderson asked his students to simplify this expression:
> $$\frac{2(a+1)}{3a} + 3 - \frac{2}{3a} - \frac{6a-2}{6}.$$

Prior to this display, the instructor had assigned, as part of the warm-up to the lesson, the simplification of this expression. There are many possible ways to simplify this expression. One is:

$$\frac{2(a+1)}{3a} + 3 - \frac{2}{3a} - \frac{6a-2}{6}$$

$$= \frac{2a}{3a} + \frac{2}{3a} + 3 - \frac{2}{3a} - \frac{6a}{6} + \frac{2}{6}$$

$$= \frac{2}{3} + 3 - a + \frac{1}{3}$$

$$= 4 - a.$$

The instructor had asked one of the prospective teachers to write down a correct simplification on the board, and then made sure that the prospective teachers understood that the original expression was algebraically equivalent to $4 - a$. Following the warm-up, the instructor discussed the Kane task (Section 3.2.2). As she brought the Kane activity to a close, she emphasized that students commonly misconstrue:

- Distributing, especially involving negatives and variables;
- Adding fractions, especially with variables; and
- Simplifying expressions of the form $ax + b$ or $a + bx$.

She then showed the slide with the truncated Anderson task and prompted the prospective teachers:

"Mr. Anderson asked his students to simplify this expression, which simplifies to $4 - a$. What are some things that could go wrong with this?"

While planning for this lesson, the instructor modified the Anderson task in response to several prior experiences using the task with preservice and novice teachers for the purpose of practicing diagnosing student work and showing a type of student error that can result from weak understanding of arithmetic operations and properties. During these instances, those working on the Anderson task would correctly identify the line in which the student work shows evidence of incorrect reasoning. However, the discussions in each of these instances had sidelined the instructor's purpose for using the task when preservice teachers or beginning instructors would question whether the records genuinely exemplified student work, often claiming that a student who had performed work shown in the mathematically correct lines would be unlikely to err as the hypothetical student had. In the instructor's prior experiences using the Anderson task, doubt as to the authenticity of the MKT task's presentation of student work had been consistently in conflict with the purpose of the task as exemplifying a kind of student error.

Regardless of the sample student work, the premise that the algebraic expression had been given to students to simplify had not previously contributed any distraction. Moreover, the types of misconceptions that could surface while simplifying this expression include those that arise in the Kane task. No matter the process used, simplifying the given expression to $4 - a$ requires understanding the distributive law (e.g., in working with the term $2(a+1)/3$ or $-(6a - 2)/6$), adding fractions involving variable expressions, and correctly simplifying variable expressions. Thus the Anderson task *as enacted* – given the contextual purpose of predicting student errors – had the potential to reinforce the instructor's purpose of identifying common student misconceptions using a mathematical frame.

Anderson task in instruction. After closing discussion on the Kane task, the instructor displayed the algebraic expression from the Anderson task and asked the prospective teachers to predict misconceptions students might display when simplifying the expression. The prospective teachers had previously simplified this expression themselves during the warm up.

"You don't distribute the negative to the $-6a$ and the 2," Tracy said.

"I think it's also tricky because there are two different denominators. When I started it, I didn't even see that there were two different denominators, so I did $3a$ but then there was a 6 and I had to go back and multiply by 2," Brittany observed, reflecting on her own solving process.

The instructor responded, "Yes, this would be a tough first problem." Brittany had put herself in the student's shoes, a key approach that serves one of the instructor's purposes, but in a way that does not support the purpose of framing student misconceptions in mathematical terms. Instead she described the work from her own point of view as though she were the student, and using language a student would likely use.

"Along the lines of the negative sign, it's hard for students to decide where the negative sign goes, so a lot of the time, they'll put it on the top and the bottom," Jason commented, taking an approach that moves back to the teacher viewpoint.

A chorus from the prospective teachers affirmed Jason's answer, "Oh yeah ... nice."

The instructor reframed, using more mathematically precise language, "So it's adding -2 divided by $-3a$." Jason agreed, but did not follow up to describe the structure of the misconception, and the conversation then shifted once more back to the prospective teachers' own work.

"I found it interesting that in Tara's work and in Danica's work, there was no finding common denominators, because you don't need to for this problem. The $2/3a$ and the $2/3a$, there are two of them, and one of them is positive, and one of them is negative, so they cancel out," Marshall observed about a possible simplification.

The instructor responded, "Which is a point about the design of the problem." Marshall's articulation provided insight into why the expression is equivalent to an algebraic expression without fractional expressions. Looking for and making use of structure such as this is a useful mathematical practice, and is part of the work of designing examples for students to use as well as understanding the mathematics. Thus Marshall was engaged in reasoning pertinent to the work of teaching; but not in a way that served the instructor's purpose of practicing mathematical articulations of student misconceptions.

"Yeah. Yeah. It's weird. So if the goal of Mr. Anderson is to get students working with common denominators, then this might not be the most useful assessment. But if that's not his goal, then it might be useful," Marshall finished. His reasoning displayed thoughtfulness about goals and problem design that served a broad purpose of the instructor for the course as a whole: cultivating in prospective teachers the norm that teaching actions should be predicated on goals. Marshall's comments demonstrate a tension between the purposes of the instructor more generally, which they serve, and moving the instruction of the moment toward the lesson's intended purpose which they do not. Tara and Danica commented on Marshall's observations.

"One of the reasons I didn't find common denominators is because I know a lot of times ... I don't remember to multiply *everything* even when I should, so for students, especially when it's a longer expression like this, where there are 4 terms, trying to multiply out to find the common denominator would cause all kinds of errors," Danica said. Danica's comment also showed reasoning that teachers engage in at times: helping students perform work accurately. Next, Tara commented, "In what could go wrong, everyone's referencing the negative, but I was thinking about combining everything at once. I wasn't sure what level," moved the instruction back to the embedded purpose of the enacted Anderson task. However, her answer displayed uncertainty about what was being asked.

The instructor replied that all grade levels of student errors are appropriate to think about; the types of misconceptions that arise in working with algebraic expressions in lower grades can persist into upper grades. Tara responded, "This isn't necessarily shown, but if you add $3a+1$ on the bottom, then they split it up and get $2/3a$ + $2/1$". She modified the problem herself to name a type of error that occurs, although the language used ("bottom", "split it up") again does not directly capture the underlying mathematical principles.

After these comments, the instructor closed the discussion about the Anderson task. Below, we examine the instructional interactions further, arguing that these interactions constitute particular affordances and tensions that were used and could have been used toward the instructor's purposes.

Affordances and Tensions in Instruction with MKT Tasks

As with any instructional resource, using MKT tasks in instruction can bring about affordances and tensions to manage toward aims of professional preparation. Our analysis suggested five such affordances and tensions that interactions between the instructor's purposes and the pedagogical context of the MKT tasks raised. In the next sections, we first provide an overview and then discuss examples of each. We then synthesize observations about the role of the elements of pedagogical context in their interaction with purposes of the instructor.

Five affordances and tensions to manage. Several affordances and tensions were present in the data and visible in the Kane and Anderson examples presented. First, the elements of pedagogical context provided an "anchor" for public reasoning about teaching, allowing prospective teachers to hear each others' reasoning in a particular teaching context. Second, the pedagogical context engaged prospective teachers. Although at times the engagement seemed to introduce ideas and framings that were not aligned with the instructor's purposes, there were also times that the engagement did serve the instructor's purposes. Third, the pedagogical context provided a structure for the instructor to use with prospective teachers as well as for the instructor to create further MKT tasks that demonstrate related work of teaching that occurs in a particular context, extending the task. Fourth, the richness of the pedagogical context contributed to unexpected or unexpectedly complex interpretations of two kinds of work: the work expected of the prospective teachers as preservice prospective teachers enrolled in a methods course, and the work of teaching as perceived by the prospective teachers when using an MKT task. We describe such interpretations as "uncharted" for the instructor in the sense that the instructor may not have been able to anticipate or steer the interpretations toward an intended purpose for the MKT task. Finally, uncertainty of purpose was visible. The prospective teachers' perception of the work of teaching can conflict with the way the task represents that work of teaching. In uncharted interpretations of work, the prospective teachers may not be aware of differences between the work that they engage in and the intended work; they engage with their interpretations unproblematically although these interpretations may be in tension with the instructor's purpose. When there is uncertainty of purpose, prospective teachers may realize that their interpretation of teaching does not cohere with the interpretation they read into the MKT task. These affordances and tensions are summarized in Table 4.

Table 4. Affordances and Tensions to manage in instruction.

Affordance	Description
Anchor example	Instructor or prospective teachers used the specifics of the pedagogical context as basis for reasoning about the work of teaching; the instructor used specifics of the pedagogical context of the MKT task itself as basis for modeling a practice.
Engaged prospective teachers toward purpose	Prospective teachers willingly reasoned within the given pedagogical context in a way that engaged them in the work of teaching. Moreover, in these instances, the prospective teachers' experience working on the MKT task supported a purpose of the instructor, even if the purpose was broader than that given by the pedagogical context or a particular planned purpose specific to that MKT task.
Structure for work	Instructor uses pedagogical context provided by MKT task, including the purpose, to structure prospective teachers' in-class activities around the work of teaching; can include extending the purpose to structure prospective teachers' in-class activities around related work of teaching not originally specified by the MKT task's context.

Source of tension	Description
Uncharted interpretations of work	Prospective teachers engage in work that differs from the work intended (by the instructor) without realizing that they are doing so. Consequently, the prospective teachers' interactions with MKT task may not be straightforward for the instructor to use toward an intended purpose, or the interactions sideline an intended purpose of the instructor.
Uncertainty of purpose	The prospective teachers' perception of the work of teaching described by the MKT task comes into conflict with the way the task represents that work of teaching.

We now discuss examples of each affordance and source of tension.

Anchor example. The pedagogical context of an MKT task can provide a common basis for public reasoning about teaching. In the Kane episode, when Marisa and Karen explain why they believe Ms. Kane's student displays some understanding, they point to the students' equating 13x – 10 with 3x. Marshall counters that it is more likely that this algebraic manipulation evidences student misconception, and goes on to point out that it happened on a "larger scale" in Ms. Kane's class. In providing concrete examples of student work, the pedagogical context was an anchor for the prospective teachers' reasoning; the context facilitated the prospective teachers' discussion by providing reasons for agreements and disagreements. They were able to hear and respond to each other in a form of professional problem solving around the specific mathematics as presented in the MKT task.

Operationally, we coded instruction with an MKT task as using the pedagogical context as an *anchor example* if the instructor or prospective teachers used the specifics of the pedagogical context as a basis for reasoning about the work of teaching, or the instructor used specifics of the pedagogical context as a basis for modeling a practice.

The reasoning facilitated by anchor examples, as well as experiences of working within the pedagogical context provided by an MKT task, can serve to engage prospective teachers toward a particular purpose of the instructor, as we discuss next.

Engaged prospective teachers toward purpose. The representation of teaching by an MKT task can potentially engage prospective teachers in ways that align with specific purposes of professional preparation. Not only are the prospective teachers' interactions with the Kane and Anderson tasks lively, there are also several instances when the prospective teachers' experiences and reasoning exemplify the purposes that the instructor had intended to convey in this lesson or across the course. Tracy's comments about the distributive property in the Anderson task and Ryan's comments about confusing surface similarities between multiplication and addition properties are examples of this. In each case, they are framing observations about potential student misconceptions in mathematical ways – one of the purposes of the instructor. When Marshall argued that the value of the algebraic problem shown in the Anderson task depends on Mr. Anderson's goals, and goes on to provide examples of goals and arguments for and against the task based on these goals, he models a disposition that the instructor had set as a broad goal for the course: that decisions in teaching should be predicated on teaching purpose.

We coded instruction with an MKT task as *engaging prospective teachers toward a purpose* if prospective teachers willingly reasoned within the given pedagogical context in a way that engaged them in the work of teaching and if the prospective teachers' experience working on the MKT task supported a purpose of the instructor. This might be an immediate purpose or one broader than that of the specific MKT task. Cases in which the potential to engage toward a purpose was only partially fulfilled are also coded, because this is an instance in which the affordance can be seen, even if it was not fully realized in this particular enactment.

The reasoning and experiences that come out of engagement with an MKT task support an instructor's use of the MKT task to structure activities for the prospective teacher in ways that help unpack the work of the profession.

Structure for work. An MKT task can be used as a resource for elaborating work of the profession that relates to but extends beyond the literal pedagogical context provided. Strictly speaking, the pedagogical context of the Kane task only included five short instances of student work responding to a fairly routine problem; a succinct description of a teacher's actions (walking around a room); and a singular purpose of determining whether or not the students' work provides evidence of correct student reasoning. However, the interactions with this context as set up by the instructor branched out into teaching actions and reasoning that are related but not directly provided by the MKT task, as well as pursuing nuances to the work of teaching directly described by the MKT task. The instructor uses the Kane task not just to provide practice for determining evidence of correct student reasoning. The instructor also structures work around the Kane task – first asking prospective teachers to "Solve the equation 5x + 8 = 13x – 10" as a warm-up, then to do the Kane task as written, then to discuss teachers' reasoning about students' reasoning as represented by the Kane task . All this serves as a potential platform for articulating the mathematical principles that would underlie correct student reasoning, and as a potential way for prospective teachers to experience a process of diagnosing student reasoning (starting from analyzing the student work for evidence of misconceptions to identifying the misconceived mathematical principles). The instructor's use of the Anderson task provides another example of using MKT tasks for structuring work; in this case, the instructor uses the Anderson task as an application of potential learning from the Kane task. She used the Kane task to situate mathematical framings that prospective teachers could then work with the Anderson task to see further examples of – and to go beyond recognizing examples of misconceptions to anticipating specific misconceptions.

We coded instruction with an MKT task as an affordance for *structuring the work* if the instructor used the pedagogical context provided by MKT task, including the purpose, to structure prospective teachers' in-class activities around the work of teaching; this can include extending the purpose to structure prospective teachers' in-class activities around related work of teaching not originally specified by the MKT task's context.

The interactions that arise within an activity's structure are shaped by prospective teachers' interpretations of the work expected of them, and this interpretation draws from prospective teachers' perception of the work of teaching.

Uncharted interpretations of work. As with instruction in general, instruction using MKT tasks can lead to interpretations of the expected work that may be in tension with an instructor's perception of how to use the MKT task towards a particular purpose. While interactions with the Kane task and Anderson task led to mathematical framings of student misconceptions, which aligned with the instructor's purpose, the interactions also led to other work that related to the work of teaching but would have needed further management to support the instructor's purpose for that lesson directly. In their interactions with the Kane and Anderson tasks respectively, Jonathan and Marshall analyzed the mathematical structure of the problem but did not make use of their observations to make inferences about specific student misconceptions of mathematical principles. In the use of the Anderson task, Brittany and Danica both discussed features of their own work that, though potentially relevant to teaching with the Anderson task, did not directly serve the instructor's purpose for using the Anderson task. These raised tensions between the prospective teachers' interpretations of the work, the purposes embedded in MKT tasks, and the instructor's purposes.

We coded instruction with an MKT task as containing *uncharted interpretations of work* when prospective teachers engaged in work that differed from the work intended (by the instructor) without realizing that they were doing so. Consequently, the prospective teachers' interactions with MKT task were not straightforward for the instructor to use toward an intended purpose, or the interactions might have sidelined an intended purpose of the instructor.

At times, the prospective teachers may attend to the purpose embedded in an MKT task in a way that is aligned with the purpose, but which raises doubt in their mind as to what the purpose means.

Uncertainty of purpose. The representation of teaching purpose in an MKT task can differ from prospective teachers' perception of teaching in ways that lead to uncertainty in how to approach an MKT task. In the Kane task, Marisa[1] and Karen perceived the Kane task's embedded purpose of evaluating student reasoning to be about student understanding, not just the validity or invalidity of reasoning. Consequently they found the binary constraint of "correct" or "incorrect" to be insufficiently nuanced. In the Anderson task, Tara understood that a purpose was to articulate potential instances of misconceptions, but she was unsure of what might count as an appropriate answer because the grade level was unspecified. Neither the originally intended nor the enacted versions of the Anderson

[1] All names are pseudonyms

task specfied the grade level, although the intended Anderson task was originally authored as part of an assessment of mathematical knowledge for teaching algebra at the middle and secondary levels. However, students' algebraic misconceptions that first arise in elementary and middle school levels can persist into their secondary education, and in this sense, the level does not need to be specified for either the enacted or intended Anderson task.

We coded instruction with an MKT task as containing *uncertainty of purpose* when the prospective teachers' perception of the work of teaching described by the task comes into conflict with the way the task represents that work of teaching.

Affordances and tensions as tensions and affordances. Interactions with an MKT task afford instruction toward a purpose when they elicit reasoning or provide experiences that align with that purpose. Interactions with an MKT task are a source of tension when they elicit reasoning or provide experiences that move away from that purpose. As stated previously, we coded instructional use of pedagogical context as an affordance or source of tension if the counterfactual absence of the element would remove the context for statements or actions that moved toward or away from a purpose. Furthermore, we noted that interactions that were a source of tension seemed to position related but distinct aspects of teaching against each other.

Affordances and tensions can, respectively, be tensions and affordances. For instance, an anchor example can engage prospective teachers in work of teaching away from a particular purpose, even if it is legitimately work of teaching.

Interactions of Elements of Pedagogical Context with Instructional Purposes

Our central question in this paper is how the features of an MKT assessment task support or limit the task's use in instruction. The prior section outlined affordances and tensions visible in instruction as we coded them in the data. Clearly many other factors mediate the use of a task in instruction besides the task's features, including the instructor's purposes but also the ways in which the instructor uses the task and manages the conversation. Our goal is not to draw conclusions about this instructor's particular use of the MKT tasks but to use it as a lens to see what such a use might reveal about the usefulness of task features. In this section we consider these affordances and tensions as they pertain to the features of the MKT tasks themseles.

Purpose provided by MKT task. Instruction structured around MKT tasks can use or extend the task's original purpose by using the purpose to orient the prospective teachers' work. For instance, in the Kane task the purpose of evaluating evidence of understanding was extended during instruction to the broader purpose of considering what a student might understand or not understand based on evidence in the student's written work, and this purpose guided the prospective teachers' work. On the other hand, if there is a conflict between teaching as represented by MKT task and prospective teachers' perception of teaching, this can create tension for the instructor to manage. For example, Marisa and Karen extend the purpose of the Kane task unprompted, possibly because for them the purpose of understanding the student's reasoning in a more nuanced way was a more reasonable purpose for Ms. Kane to have than simply deciding if the reasoning was valid or not. As it turns out, this extension was in a direction intended by the instructor, allowing her to capitalize on what might otherwise have been a problematic misunderstanding of what the Kane task was asking for.

One purpose of each original MKT task is to orient the test taker in to how to use the task. Used in instruction, the purposes given explicitly to the prospective teachers by the instructor also oriented the prospective teachers work, but in a less constrained way. The purpose of the original MKT task is clearly constrained in order to support its use in assessment – a clear and defensible answer depends on a clear statement of what is to be evaluated. But this feature, key for assessment purposes, may be inauthentic to teaching practice in ways that create tension when the task is used in instruction. In our data, the instructor's adjustment to a more open-ended purpose was productive, but also added instructional complexity for the instructor to manage, and the potential for the prospective teachers to become confused about the purpose.

Another way in which purpose was in tension with instruction is when the prospective teacher's perception of teaching was in conflict with the task's representation of teaching. For example, the Kane task, by asking the question "Does provide"/"Does not provide", implicitly sets this as a reasonable criterion on which to judge student work. Prospective teachers working on this task sometimes did not know how to answer the question because they believed that each of the sample pieces of work potentially showed some understanding. In a sense, attending to all evidence of student understanding, even when it is weak evidence or evidence of partial understanding, is a more authentic and desirable way to interact with student work than the interaction assumed by the original task. Again, because the MKT task was originally constructed for assessment purposes, it requires a clear task statement. It may seem obvious to say that the closer the instructor's purpose was to the purpose of the intended MKT task, the more likely it was that this feature would support use in instruction. However, one can imagine this happening with any task that prompts the test taker similarly, and we argue that this is potentially a point on which it is likely tensions will arise between teacher education purposes and the original assessment purposes for which many MKT tasks were written.

Records of practice. Records of practice (both the given mathematical problem and the student work samples around that problem) engaged prospective teachers in ways that provided relevant experience that promoted conviction toward purposes of the instructor. They provided reference points for prospective teachers to hear and respond to each other's contributions. On the other hand, they were also a source of complexity in instruction in the way that prospective teachers' engaged with them, as can be seen in the work prospective teachers engaged in around the Kane task. Prospective teachers might meta-analyze the problem in ways that do involve interesting MKT but that don't have to do with the point at hand, as for example in Jonathan's analysis of design of records in Kane. Such records can also serve to "set the tone" for further conversation. For example, discussion about the Kane task used student-like language and the instructor was unable to bring more mathematically precise descriptions to bear; the use of student work to ground the conversation may have set the tone that prospective teachers were discussing things as students would, creating tension with the instructor's purpose of moving the discussion away from student-talk.

Records of practice provide anchor examples in that prospective teachers and the instructor can base arguments on them and for the instructor to use in modeling practice. In the Anderson task, for example, Marshall analyzed the task for information that it may or may not provide about students' understanding, and hypothesized about Mr. Anderson's purpose based on the specifics given, creating potential for a more general discussion about teaching purpose within the very specific examples given.

For records of practice to be effective reference points, they need to be believable. This was a motivation, for example, for the instructor's modification of the Anderson task based on past use of the task in which prospective teachers did not believe the represented student error to be consistent with the level of the student's work as represented in other steps. This is a particularly important point of tension to note because preservice teachers do not necessarily have the experience needed to correctly determine if records of practice are authentic, particularly when they represent student work, and this is often precisely why an instructor might want to introduce such examples. In the Kane task, for example, the instructor was able to use the task to help prospective teachers articulate a particular and common student error (combining like terms), and part of doing so was showing that students do in fact commit such errors. The primary mechanism for doing this was through the records of practice. Had the prospective teachers found the representations to be inauthentic, however, as they did in past use of the Anderson task, they might instead have simply dismissed the work as something student don't do and been unable to engage. Authenticity is in part a feature of the given MKT task – often there is research-based evidence that supports claims that particular student errors are or are not, in fact, common errors – but authenticity is also in part in the eye of the beholder.

Student background and organization. Neither of the example tasks provided in this paper provided detail about the level of the students in the task scenario or the class organization; the information given contributes to the plausibility of the task. For example, in the Kane task it is not strictly necessary for solving the task to know that the teacher is walking around the classroom, but this image creates a context in which the question "valid or not" is a reasonable question to be asking. These features were, however, important in the instructor's extension of the tasks to new purposes, as, for example, in the Anderson task, where considering the level of the students became a part of the conversation about anticipating what kinds of mistakes a student might be likely to make.

Discussion: Future Directions for Professional Preparation Using MKT Tasks

This work begins to bridge the research and practice gap between MKT assessment and MKT as curriculum for teacher education, and contributes to the conceptualization of MKT as mathematical knowledge integrated with teaching purpose. The reported study provides examples of MKT assessment tasks used as a resource for MKT instruction, which has significance both theoretically and practically. Prior research shows that the integration of teaching purpose is essential for valid assessment; this study further shows that this integration can be used to communicate MKT as professional knowledge because it can be used to problematize mathematical aspects of teaching. Such examples are of direct use to instructors in considering whether and how to make use of MKT assessment tasks now becoming publicly available, such as the released MET measures (Bill and Melinda Gates Foundation, 2012). Further, there is evidence of interest within the professional community from instructors who wish to use these assessment-oriented materials in instruction.

Positing that instruction with MKT tasks is shaped by interactions between purpose and pedagogical context, next steps for research and practice include mapping the instructional *geography of purposes* that can be interpreted from MKT tasks, developing norms and processes for designing MKT tasks for use in instruction, and improving practitioners' abilities to manage instructional interactions towards particular purposes. These next steps can leverage current directions in the field of mathematics education and teacher education.

The uses made of the Kane and Anderson Tasks in this study suggest that multiple purposes can be interpreted from and layered onto the context provided by an MKT task, by both instructors and prospective teachers. In general, whether students are K-12 students or prospective teachers, and the teacher is an instructor or a K-12 teacher, there are many purposes that a teacher may choose to use a particular activity toward. Some of these purposes may turn out to be better suited to the intended activity than others. As well, students may read purposes that are different from the purposes of the teacher into the intended activity. There may be an instructional *geography of purposes* – that, given a particular purpose, some purposes are more likely to pair well with the given purpose than others, and that some purposes are more likely to be misread into a given purpose than others. Such purposes would be located nearer to the given purpose than others in the geography, which would be located farther away. For instance, in terms of purposes of the instructor to provide to prospective teachers, it seems plausible that the purpose "determine whether the work shown below provides evidence of correct reasoning" would be near the purpose "diagnose the misconceptions that may be likely given this work", and these purposes might be near the purpose "select or design an example that would elicit a this misconception". We believe that if this *geography of purposes* is to be useful, it must be closely aligned with current efforts to decompose teaching practice into teachable and learnable components (e.g., Grossman et al., 2009; Ball & Forzani, 2009), as well as efforts to understand the work of teaching that arises when teaching particular subjects (e.g., Herbst et al., 2009).

Designing MKT tasks for instruction brings with it, at a minimum, the challenges of developing an initial concept of the pedagogical context, revising the task to clarify the features of the pedagogical context and how they function toward a solution to the MKT task that is true to the work of mathematics teaching, and ensuring that the alignment of features and mathematics teaching remains through revisions. Writing valid MKT assessment tasks is a time-consuming process; one expert item writer estimates the total time required to produce a usable MKT assessment task at around 10.5 hours (B. Weren, personal communication, January 10, 2014). It makes sense to leverage the time and energy put into design for multiple uses, even where those uses may require adjustment of the task itself. Recent efforts at writing MKT assessment tasks show promise that the process can become less time intensive by building on efforts of previous MKT task writing processes (Herbst & Kosko, 2012). As well, in the area of mathematics problem writing, processes for developing mathematics problem writing skills are being developed for tasks exemplifying the Common Core content standards, and eventually, practice standards (e.g., Illustrative Mathematics, n.d.).These efforts could potentially be leveraged to develop the professional skill of mathematics instructors in writing or adapting MKT tasks that can be shared and used across teacher education programs.

Fluency in the *geography of purposes* and in constructing MKT tasks both contribute to improving instructors' ability to manage instruction with MKT tasks. The *geography of purposes* could help instructors mediate how prospective teachers may interpret purposes, design modules for prospective teacher preparation and development, and structure curricula for teacher education. By having a better grasp of nearby purposes that prospective teachers may read into a given purpose, instructors may be better able to predict how prospective teachers may interpret an MKT task, observe prospective teachers' interpretations more readily, and use prospective teachers' interpretations to help reinforce or craft learning goals for prospective teachers.

A set of MKT assessment tasks has been recently released by the Measures of Effective Teaching (Bill and Melinda Gates Foundation, 2012), and other validated assessments are being developed; in time, these efforts may also release sample tasks. These validated tasks constitute a dominant representation of MKT and its use to educators. A fluency in the *geography of purposes* would benefit instructors who want to build professional preparation activities from released MKT assessment tasks, or of any other MKT tasks that instructors may come across or design. Understanding the process of constructing MKT tasks may help instructors adapt existing MKT tasks to local contexts. Beyond understanding the terrain purposes and the construction of MKT tasks, there is also a need to develop pedagogical knowledge for using MKT tasks, such as frameworks for orchestrating discussion.

Although MKT tasks can convey aspects of the teaching profession and its entailed knowledge and practices, the validation of MKT tasks with practicing teachers means that there will be inherent mismatches between MKT tasks as assessment and MKT tasks as problems to use in the instruction of preservice teachers. This may be especially important where records of practice are presented by an MKT task and where the instructor's purpose is to use these records as examples or evidence of what student work tends to look like. The study further suggests that further supports are needed for instructors' effective use of MKT tasks in instructional tasks. Our findings indicate both promise and caution in using MKT assessments as resources for teacher education.

References

Ball, D.L., & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education, 60*(5), 497-511.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407. Retrieved from http://dx.doi.org/10.1177/0022487108324554

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A.,… Tsai, Y. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180. Retrieved from http:// dx.doi.org/10.3102/0002831209345157

Bill and Melinda Gates Foundation. (2012). *Measures of effective teaching: Item-level instrument files, 2009-2011, ICPSR34345-v2.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research. Retrieved from http://dx.doi.org/10.3886/ICPSR34345.v2

Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des mathematiques 1970-1990.* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. & Trans.) Dordrecht, The Netherlands: Kluwer.

Cohen, D. K., Raudenbush, S., & Ball, D. L. (2002). Resources, instruction, and research. In F. Mosteller & R. Boruch (Eds.), *Evidence matters: Randomized trials in education research* (pp.80-119). Washington, DC: Brookings Institution Press. Retrieved from http://dx.doi.org/10.1029/2001WR000300

Gitomer, D., Phelps, G., Weren, B. Howell, H., & Croft, A. (in press). *Evidence on the validity of content knowledge for teaching assessments*. In T. Kane, K. Kerr, & B. Pianta (Eds.), *MET Research Volume*.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055-2100.

Herbst, P., Chen, C., Weiss, M., & González, G., with Nachlieli, T., Hamlin, M., & Brach, C. (2009). "Doing proofs" in geometry classrooms. In M. Blanton, D. Stylianou, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 250-268). New York, NY: Routledge.

Herbst, P. & Kosko, K. (2012). Mathematical knowledge for teaching high school geometry. Paper presented at the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Kalamazoo, MI: PME-NA.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371-404). Charlotte, NC: Information Age Publishers.

Hill, H. C., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371- 406. Retrieved from http:// dx.doi.org/10.3102/00028312042002371

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, 105, 11-30. Retrieved from http://dx.doi.org/10.1086/428763

Illustrative Mathematics (n.d.). *Summary of reviewer training objectives.* Retrieved from http://www.illustrativemathematics.org/summary_tm_objectives

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (pp. 17-64). New York, NY:

American Council on Education & Praeger.
Lai, Y., Jacobson, E., & Thames, M. (2013). *The role of pedagogical context in measures of specialized and pedagogical content knowledge*. Paper presented at the meeting of the American Educational Research Association. Retrieved from the AERA Online Paper Repository.
Learning Mathematics for Teaching. (2008). *Mathematical Knowledge for Teaching (MKT) Measures: Mathematics Released Items*. Author: Ann Arbor, MI.
National Academy of Sciences. (2010). *Preparing teachers: Building evidence for sound policy* (pp. 67-68). Washington, DC: The National Academies Press.
Phelps, G., Howell, H., & Kirui, D**.** (2013). *Assessing content knowledge for teaching: Situating knowledge assessment in teaching practice*. Manuscript in preparation.
Phelps, G., Howell, H., & Schilling, S. (2013). *Assessing the development of mathematical knowledge for teaching*. Manuscript in preparation.
Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, 6(1), 43-74. Retrieved from http://dx.doi.org/10.1162/EDFP_a_00022
Sleep, L. (2012). The work of steering instruction toward the mathematical point: A decomposition of teaching practice. *American Educational Research Journal, 49*(5), 935-970.
Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14. http://dx.doi.org/10.3102/0013189X015002004
Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, *33*(2), 455–488. Retrieved from http://dx.doi.org/10.3102/00028312033002455

Characteristics of Successful Programs in College Calculus: How Calculus Instructors Talk about their Students

Sean Larsen	Estrella Johnson	Dov Zazkis
Portland State University	Virginia Tech	Rutgers

Abstract

The CSPCC (Characteristics of Successful Programs in College Calculus) project is a large empirical study, investigating mainstream Calculus 1, that aims to identify the factors that contribute to successful programs. The CSPCC project consists of two phases. Phase 1 entailed large-scale surveys of a stratified random sample of college Calculus 1 classes across the United States. Phase 2 involves explanatory case study research into programs that were identified as successful based in part on the results of the Phase 1 survey. During our case study site visits, we interviewed calculus instructors and asked a number of questions that prompted them to discuss their students. ***The purpose of the analyses we will present here is to characterize the ways that calculus instructors talk about their students***. To do so, we will examine instructor survey responses and analyze instructor interviews conducted at the case-study institutions (PhD and Bachelors granting levels).

Study Background and Research Questions

The CSPCC (Characteristics of Successful Programs in College Calculus) project is a large empirical study, investigating mainstream Calculus 1, that aims to identify the factors that contribute to successful programs. The CSPCC project consists of two phases. Phase 1 entailed large-scale surveys of a stratified random sample of college Calculus 1 classes across the United States. Phase 2 involves explanatory case study research into programs that were identified as successful based in part on the results of the Phase 1 survey. Specifically, institutions were selected based on student *persistence* (continuing on to take Calculus 2), *success* (pass rates in Calculus 1), and reported increases in students' *interest*, *confidence*, and *enjoyment* of mathematics as a result of taking Calculus 1. This second phase will lead to the development of a theoretical framework for understanding how to build a successful program in calculus and in illustrative case studies for widespread dissemination.

During our site visits, we interviewed calculus instructors and asked a number of questions that directly or indirectly prompted them to discuss their students. ***The purpose of the analyses we will present in this preliminary report is to characterize the ways that calculus instructors talk about their students in an attempt to understand how their perceptions of their students may be related to their approach to teaching calculus.*** To do so, we will examine instructor survey responses for all surveyed institutions (of all types) and analyze instructor interviews conducted at the selected case-study institutions at the PhD and Bachelors granting levels.

Relation Research literature & Theoretical Perspective

Our decision to focus on instructors' views of their students is informed by a belief that teachers' views of students inform their classroom interactions and teaching practices. This notion is supported by Blumer's (1969) perspective of Symbolic Interactionism (SI). The first premise of

SI states that, “human beings act toward things on the basis of the meanings that the things have for them” (Blumer, 1969, p. 2). From this perspective, teachers act towards students (at least in part) based on the meanings that students have for them. For example, if teachers view students as disengaged and incapable, this may influence their instructional decisions. This idea is supported by empirical research looking at relationships between teachers’ beliefs and their teaching practice. For instance, Thompson (1984) presented a case study of a middle school mathematics teacher, Lynn. As described by Thompson, the most important factor influencing Lynn’s teaching practice was “her low expectations of the students and her pervasive concern to get through the day’s lesson in a manner that would minimize the potential for student disruptive behavior” (p. 117). Thompson goes on to state that, “implicit in her [Lynn’s] attitude was a belief that little could be accomplished in terms of teaching and learning given the poor disposition of the students and the wide diversity in their background knowledge” (p. 117).

Research methods

Our research team conducted site visits at five bachelors granting institutions and five PhD granting institutions (Table 1). While on campus, we interviewed students, instructors, administrators, and others involved in the calculus program at the institution. This report will focus on the instructor interviews. We interviewed a total of 54 instructors over the course of the 10 case study site visits.

Table 1: Institution characteristics

Institution	Highest Degree	Type	Unduplicated head count	Calculus Class size	Instructor Interviews
B1	Bachelors	Non-Profit	9,000	25-30	7
B2	Bachelors	Public	30,000	30-40	3
B3	Bachelors	Non-Profit	4,500	20-25	6
B4	Bachelors	Non-Profit	4,000	20-28	3
B5	Bachelors	Non-Profit	6,000	25-30	6
P1	PhD	Non-Profit	40,000	30-200	5
P2	PhD	Public	7,500	35-45	5
P3	PhD	Public	31,000	200-250	6
P4	PhD	Public	31,000	30-35	8
P5	PhD	Non-Profit	6,000	30-35	5

The first stage of our data analysis (of interviews) involved “tagging” all transcripts by identifying the relevant topics addressed by the interviewee in each of their responses. For example if an instructor said, “some of my students are not prepared to handle the algebraic procedures needed to use the derivative concept on the application problems that we put on our common exams” this response would be tagged with the following codes: *student subject characteristic*, *assignments and assessments*, *course coordination*, and *content*. (The scheme consists of 24 tags.) Two coders tagged each transcript independently and each response was ultimately tagged with the union of tags applied by the two coders. Using these tagged transcripts in HyperResearch, we will run reports of all instructor responses that refer to *student characteristics.* We will also run reports of to identify which of these responses also include statements about *teaching and learning*, *pacing*, *assignments* and *assessments*, or other topics that will allow us to contextualize the instructors' statements about students and examine relationships between the instructors' views about students and their (stated) instructional practices. We will then analyze instructors' comments about their students in order to identify categories of meanings that students have for teachers. Our goal is to develop a taxonomy of categories that we can then coordinate with our analyses of the teachers' instructional practices in order to generate conjectures about relationships.

Results of the research

Survey Results. There were several questions posed on the instructor pre and post surveys that relate to characteristics of students:

- (Pre) Approximately what percentage of students currently enrolled in your Calc I course do you expect are academically prepared for the course?
- (Pre) Estimate the percentage of students currently enrolled in Calc I that will: (*outcome options*)
- (Post) Approximately what percentage of your students were prepared for the course
- (Post) What percentage of students received a C or better?
- (Post) All students in beginning calculus are capable of understanding the ideas of calculus (*Likert scale*)

Taken as a group, the instructors at selected case study schools (all levels) predicted higher percentages of prepared students and higher percentage of successful students on the pre-surveys. Further, these instructors reported higher levels of preparation and success on post-surveys. All of these differences were statistically significant. These instructors also expressed a higher level of agreement that students in beginning calculus are capable of understanding the ideas of calculus. However, this difference was only significant at a .1 level. Of course, the differences we see in predicted/reported levels of preparation and success could simply indicate that these teachers tend to work with students who are better prepared and more successful. Therefore, given the survey data alone, it is impossible to draw any conclusions about the impact of the instructors' views of their students on their practice (and their students' success). In an effort to unearth additional insight into how calculus instructors view their students at the selected case study schools, and to better understand what impact this may have on their practice, we have begun the process of carefully analyzing our instructor interviews.

Sample of Interview Results and *Directions for Ongoing Analysis*

Analysis of the interviews is just underway, but it is already clear that calculus instructors talk about students in a variety of ways and that there are relationships between these views and their (reported) instructional practices. Below we see three different quotes (all from instructors at institution B3.

Dr. Young: Yeah, they're not good at algebra. It's getting worse and worse. Yeah. They need an algebra class, except they've had like 14 algebra classes before now.

Dr. Adams: I think -- well, I like the age that I teach. They tend to be freshmen. So I get to work with students who are in the process of transitioning to college life.

Dr. Bell: …when we have so many students coming in with AP Calculus thinking they understand calculus and -- well, my experience has been is that they know how to do problems mechanically, but their understanding is not very deep. And so that makes it very hard because they have this "I already know what you're going to do" like sort of mentality, and "You can't teach me anything." And that is a really hard thing to overcome.

From these quotes, we can see that instructors may think about students in terms of their academic abilities, their attitudes, and where they are in their lives. Ongoing analyses will enable us to add to and refine these categories as well as explore relationships to teaching practice.

Questions for the Audience:

What data might we look at in order to better understand our quantitative analysis of the survey questions (e.g., student SAT scores)?

How could we use our student focus group interviews to address our research question?

References

Blumer, H. (1969) *Symbolic Interactionism; Perspective and Method*. Englewood Cliffs, NJ: Prentice-Hall

Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, *15*(2), 105–127. doi:10.1007/BF00305892

IMPLEMENTING INQUIRY-ORIENTED INSTRUCTIONAL MATERIALS IN UNDERGRADUATE MATHEMATICS

Christine Larson Florida State University	Megan Wawro Virginia Tech	Michelle Zandieh Arizona State University
Chris Rasmussen San Diego State University	David Plaxco Virginia Tech	Katherine Czeranko Arizona State University

Over the past years, research in the RUME community has driven the development of inquiry-oriented instructional materials in a number of undergraduate mathematics content areas including abstract algebra, differential equations, and linear algebra. Literature at the K-12 level has documented challenges inherent to scaling up the implementation of this kind of instruction. In this study, we explore how instructors make sense of and implement inquiry-oriented instructional materials in undergraduate mathematics, and the nature of supports these instructors report using and wanting when planning for instruction. We consider instructors' interpretations and desired supports as they relate to prior pedagogical experience and institutional setting. Data is taken from surveys, interviews, and video-taped instruction of three participating instructors at three different institutions as they work to implement two inquiry-oriented instructional units in undergraduate linear algebra.

Key Words: Linear Algebra, Inquiry-Oriented Instruction, Scaling Up

Over the past years, research in the RUME community has driven the development of inquiry-oriented instructional materials in a number of undergraduate mathematics content areas including calculus, abstract algebra, differential equations, and linear algebra (e.g., Swinyard, 2011; Larsen, 2009; Rasmussen & Kwon, 2007; Wawro, Rasmussen, Zandieh, Sweeney, and Larson, 2012). Literature at the K-12 level has documented challenges inherent to scaling up the implementation of this kind of instruction, particularly considering the role of institutional factors (e.g., Elmore, 2004; Gamoran et al., 2003; Bond, Boyd, & Montgomery, 1999). In this study, which is part of a newly funded NSF grant, our project team explores how instructors make sense of and implement two inquiry-oriented instructional sequences in undergraduate linear algebra. Additionally, we identify supports these instructors report using and wanting when planning for instruction, and we consider instructors' interpretations and desired supports as they relate to prior pedagogical experience and institutional setting.

This project will enable us to contribute to a growing body of literature in undergraduate mathematics education that considers issues of instructor learning and practice. Specifically, we will consider literature from educational policy and leadership, teacher professional development, instructional practice, and instructional design. In doing so, we hope to move toward the development of a framework for sharing and scaling up inquiry-oriented instruction in undergraduate mathematics.

This preliminary report targets two broad goals. First, we aim to work with participating instructors to inform the design of the instructional materials themselves. This will allow us not only to incorporate instructor feedback to refine the materials, but also to better understand the ways in which instructors interpret and use these kinds of materials to plan for instruction. Second, we aim to work with participating instructors to better understand the kinds of supports that will be needed by future instructors in order to implement the materials as intended. As such, we aim to answer the following research questions:

1. How do instructors make sense of inquiry-oriented instructional materials in undergraduate mathematics?
2. What forms of support (including material supports and professional interactions) do instructors report using, needing, and wanting to accompany these instructional materials?
3. In what ways do instructors' prior pedagogical experiences (e.g. training, mentorship, teaching experience) and institutional settings (e.g. formal structures of their university/college/school/department, departmental culture, informal colleague support) relate to instructors' interpretation and implementation of such materials?

Literature

Effectively implemented inquiry-oriented instructional approaches have been related to improved levels of conceptual understanding and equivalent levels of computational performance in areas ranging from K-12 mathematics, to undergraduate mathematics, physics, and chemistry (e.g., Rasmussen & Kwon, 2007; Cai, Nie, & Moyer, 2010; Deslauriers, Schelew, & Wieman, 2011; Lewis & Lewis, 2005). More specifically, students who engage in cognitively demanding mathematical tasks have shown greater learning gains than those who do not (Stein & Lane, 2006). Furthermore, Stein and Lane (2006) found that those gains were greater in classrooms in which students were encouraged to use multiple representations, multiple solution paths, and where multiple explanations were considered; gains were lower in classrooms where the teacher demonstrated a process students could use to solve the task.

Research has shown, however, that instructors often struggle to transition to an inquiry-oriented teaching style. For instance, Wagner, Speer, and Rossa (2007) investigated the struggles that Rossa, a university mathematician with little inquiry-oriented teaching experience, had in trying to implement an inquiry-oriented approach to differential equations. These challenges included facilitating productive whole class discussions, identifying appropriate learning goals, and assessing students' progress toward these goals (e.g., making decisions about what constitutes an adequate student understanding and how/when to move on to a new topic). A central theme in this and other research documenting mathematicians' efforts to implement inquiry-oriented curricula is the need for an understanding of student thinking to plan for and lead discussions that effectively build on students' solution strategies (Johnson & Larsen, 2012; Speer & Wagner, 2009).

Inquiry-oriented Instructional Materials

This proposal is part of a larger study that aims to produce (a) student materials composed of challenging and coherent task sequences that facilitate an inquiry-oriented approach to the teaching and learning of linear algebra; (b) instructional support materials for implementing the student materials. Our framework for designing student materials draws on heuristics of Realistic Mathematics Education (RME; summarized by Cobb, 2011). First, a task sequence should be based on experientially real starting points. In other words, the initial tasks of a sequence should be set in a context that enables students to engage in, interpret, and make some initial mathematical progress. Second, the task sequence should be designed to support students in making progress toward a set of associated mathematical learning goals. Third, classroom activity should be structured so as to support students in developing models-of their mathematical activity that can then be used as models-for subsequent mathematical activity.

Finally, with instructor guidance, students' activity evolves toward the reinvention of formal notions and ways of reasoning about the mathematics initially investigated. This framework facilitates a transition from students' current, informal ways of reasoning about key concepts in linear algebra towards more formal, mathematically sophisticated ways of reasoning.

Following the model described by Lockwood, Johnson & Larsen (2013), the instructor support materials for each instructional unit includes rationale for the instructional sequence, insight into student thinking, and ideas about implementing the unit. The rationale for these features is firmly rooted in the literature (Ball & Cohen, 1996; Collopy, 2003; Davis & Krajcik; 2005) and has been shown to be useful for instructors (Lockwood et. al., 2013).

Participants & Data Sources

In this project, we work with three participating instructors at three different institutions to implement two instructional units in an intact introductory linear algebra class. Each participating instructor is paired with a member of our project team who was involved in the development of the instructional materials. The purpose of this partnership is to provide support for the participating instructor as they plan for instruction with the materials and to obtain feedback on planning with and implementing the materials. In addition, we interview and survey each participating instructor to document the setting in which they teach, their pedagogical experiences and training, the student populations they serve, and their views of teaching and linear algebra. As part of our efforts to develop instructional support materials, we collect information on both how instructors make use of those support materials and additional supports the instructors feel would be helpful. This will allow us to embed research-based revisions of the instructional support materials into the ongoing work of the project.

Data collection for the first phase of the project is currently ongoing (to be completed at the end of the Fall 2013 semester), and the data set will consist of three primary components. First, each participating instructor will work with his/her project team member to co-plan each instructional unit using the instructional materials. This audio-taped co-planning interview will document the instructor's learning goals for his/her students, the activities he/she intends to use with the students (noting any modifications and/or supplementation of the provided materials), and how the instructor intends instruction of the unit. If the instructor develops a written lesson plan, the cooperating project team member will ask to keep a copy. Second, the research team will videotape the instructor participants' enactment of the units. The camera will be focused on the instructor, rather than the students, during this data collection. This will allow the research team access to data (other than self-report data from the instructors) to examine difficulties and successes within the lessons. Third, the participating instructors will debrief with their project team member. This debriefing will be audio recorded. The project member will ask the instructor to reflect on the implementation of the unit and will follow a debriefing protocol to specifically probe on adjustments made to the plan in teaching the unit and the rationale for those adjustments. This approach allows us to collect data on instructors' implementation of the units while also providing them with individualized instructional support.

Methods of Analysis

To answer our first two research questions about instructor interpretation of materials and forms of support, we will perform a grounded analysis (Strauss & Corbin, 1998) of audio data from the co-planning and debriefing interviews as well as video of their classroom implementation of the units. Here we will consider the participating instructors' feedback on the instructional materials, their description of their plans for implementing them, and the questions they asked and advice they received from their supporting project team member. We will triangulate this with their video-taped instruction to for additional data about their implementation, specifically examining how they facilitated whole class discussion before and after students worked on the tasks, the use and introduction of language and notation, and key ideas emphasized. This will help us document and understand the variety of ways in which instructors may go about implementing the tasks in their classrooms and modifications they might make to the instructional sequence.

We will rely on comparative case study methodology (Yin, 2003) for our analysis to theorize about relationships among instructor background, institutional characteristics, desired support, and material interpretation/implementation. This will draw on data from the aforementioned interviews, survey, and video-taped instruction. For this work, we will develop a matrix to analyze key dimensions that we believe to be important contextual factors that might influence participants' instructional decisions. These include factors such as the nature of the introductory linear algebra course at their institution, the student population they serve, and, instructor's pedagogical training and experience, departmental culture. We can then consider how particular ways of interpreting, modifying, and implementing the materials align with these characteristics as well as the forms of instructional supports desired and received from their partnering project team member.

Questions for Audience Discussion

- A future goal is to explore what might be meant by "effective implementation" of the instructional materials. How might this be defined and measured in such a way that honors instructors' expertise as well as the intent of the materials?
- What institutional factors might afford or constrain instructor efforts to learn about and implement instructional innovations that we have not yet considered here?

References

Ball, D. L. & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instruction reform? Educational Researcher, 25(9), 6-8, 14.

Bond, S. L., Boyd, S. E., & Montgomery, D. L. (1999). *Coordinating resources to support standards-based mathematics education programs*. Chapel Hill, NC: Horizon Research Inc.

Cai, J., Nie, B., & Moyer, J. C. (2010). The teaching of equation solving: Approaches in Standards-based and traditional curricula in the United States. *Pedagogies: An International Journal*, *5*(3), 170-186.

Cobb. P. (2011). Learning from distributed theories of intelligence. In E. Yackel, K. Gravemeijer, & A. Sfard (Eds.), A journey into mathematics education research: Insights from the work of Paul Cobb (pp. 85-105). New York: Springer.

Collopy, R. (2003). Curriculum materials as a professional development tool: How a mathematics textbook affected two teachers' learning. The Elementary School Journal, 103(3), 287-311.

Davis, E. A. & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. Educational Researcher, 34(3), 3-14.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in large-enrollment physics class. Science, 332(6031), 862-864.
Elmore, R. F. (2004). *School reform from the inside out*. Cambridge, MA: Harvard Education Press.
Gamoran, A., Anderson, C. W., Quiroz, P. A., Secada, W. G., Williams, T., & Ashman, S. (2003). *Transforming teaching in math and science: How schools and districts can support change*. New York: Teachers College Press.
Johnson, E., & Larsen, S. (2012). Teacher listening: The role of knowledge of content and students. Journal of Mathematical Behavior, 31, 117-129.
Larsen, S. (2009). Reinventing the concepts of group and isomorphism. Journal of Mathematical Behavior, 28, 119-137.
Lewis, S. E., & Lewis, J. E. (2005). Departing from lectures: An evaluation of a peer-led guided inquiry alternative. Journal of Chemistry Education, 82, 135-139.
Lockwood, E., Johnson, E., & Larsen, S. (2013). Developing instructor support materials for an inquiry-oriented curriculum. *The Journal of Mathematical Behavior*.
Rasmussen, C., & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *The Journal of Mathematical Behavior*, *26*(3), 189-194.
Speer, N. M., & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. Journal for Research in Mathematics Education, 40(5), 530–562.
Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. Educational Research and Evaluation, 2, 50-80.
Strauss, A. L., & Corbin, J. M. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). Thousand Oaks: Sage Publications.
Swinyard, C. (2011). Reinventing the Formal Definition of Limit: The Case of Amy & Mike. Journal of Mathematical Behavior, 30, 93-114.
Wagner, J., Speer, N. M., Rossa, B. (2007). Beyond mathematical content knowledge: A mathematician's knowledge needed for teaching an inquiry-oriented differential equations course. Journal of Mathematical Behavior, 26, 247-266.
Wawro, M., Rasmussen, C., Zandieh, M., Sweeney, G. F., & Larson, C. (2012). An Inquiry-Oriented Approach to Span and Linear Independence: The Case of the Magic Carpet Ride Sequence. *PRIMUS*, *22*(8), 577-599.
Yin, R.K. (2003). *Case study research: Design and methods (3rd ed.).* Thousand Oaks, CA: Sage.

85

CALCULUS STUDENTS' EARLY CONCEPT IMAGES OF TANGENT LINES

Renee LaRue, Brittany Vincent, Vicki Sealey, Nicole Engelke
West Virginia University

In this study, we use Newton's method as a means to examine first-semester calculus students' understanding of tangent lines. Within that context, we found that many students had difficulty sketching and describing tangent lines. We examined the language students use to describe tangent lines as well as their graphical illustrations of tangent lines. Task-based interviews were conducted with twelve first-semester calculus students who were asked to verbally describe a tangent line, sketch tangent lines for multiple curves, and use tangent lines within the context of Newton's method. Six prominent categories describing students' concept images of tangent lines emerged, and we found that individual students often possessed multiple concept images. Furthermore, data shows that these concept images were often conflicting, and students were usually willing to modify their concept images in different contexts.

Keywords: Calculus, Tangent Line, Newton's Method, Derivative

Introduction

In this study, we examine calculus students' concept images of tangent lines and how these concept images manifest as students use tangent lines to graphically understand the iterative process of Newton's method. We compare students' verbal and graphical responses when asked "What is a tangent line?" with their usage of tangent lines within the context of Newton's method. We found that many students were unable to sketch tangent lines even remotely resembling the correct line, making it impossible for them to truly understand how Newton's method works. In some cases, we found that students were willing to modify their understanding of tangent lines to fit what they believed Newton's method was trying to do.

Existing research about students' understanding of tangent lines often focused on the connection between lines tangent to a circle and lines tangent to a curve (e.g. Biza and Zachariades (2010), Páez Murillo and Vivier (2013), Vinner (1982)). Vinner (1982) states that most educators use one of two approaches when introducing lines tangent to a curve in calculus. The first approach is to assume that the students will intuitively develop the correct understanding based on previous knowledge from geometry and will need no specific instruction. The second uses the geometrical approach and defines a tangent line as the result of a limiting process of secant lines. An instructor using this approach would later make the connection between this limiting process and the limiting process used in the definition of the derivative.

According to Biza and Zachariades (2010), there are two main viewpoints from which to approach the study of tangent lines – analysis and geometry. In analysis, "the existence of tangent line at a point is a property of the curve at this point," which they consider to be a local perspective (p. 219). In geometry, the tangent line is a property of the line together with the entire curve, which the authors consider to be a global perspective. Additionally, they state that in order to properly understand tangent lines in any context, it is necessary to hold both perspectives at the same time. Biza (2008) discussed these viewpoints when she described several specific concepts that may contribute to students' misconceptions about tangent lines. Most notable for our research are her observations that students often believe that tangent lines must have only one point in common with the curve and that, when extended far beyond the point of tangency, the line could not intersect the curve. Biza summarized these ideas, saying that certain properties of tangent lines that arise in *specific* contexts were taken by the students to be defining characteristics of *all* tangent lines.

Theoretical Perspective

For the design and analysis of this study, we refer to Piaget's structuralism (1970, 1975), which allows us to look at structures as a whole as well as the individual parts that make up the structure. For this study, we focus on the individual parts (e.g. tangent lines, intercepts, etc.) that make up the structure for Newton's method. The interview questions were targeted at addressing these parts in order to discover the strengths and weaknesses of these pre-existing component structures as well as to determine how these parts help students to create an understanding of Newton's method. According to structuralism, students need to be engaged in activities that allow them to *do* something with tangent lines that is reflected in and regulated by the underlying structure of its component parts. It is then that students can reflectively abstract to further develop that structure as well as strengthen connections between existing structures.

Piaget uses the terms assimilation and accommodation to describe how new information is incorporated into structures that have been previously established. When taking in new information, it will either fit into the previously established structure or it will cause conflict. If the new information does not fit, the learner must either adjust the new information to fit with the existing structure via assimilation, or he/she must reorganize the existing structure to allow for the new information via accommodation. It is important to note that assimilation and accommodation occur simultaneously (Piaget, 1970). In regards to this research, the materials used encouraged students to incorporate the structure of Newton's method into their existing tangent line framework (via assimilation), while also allowing students to modify aspects of their tangent line structures (via accommodation). Through this process, the students make connections between their understanding of both tangent line and Newton's method.

Offering additional insight for our analysis is the work of Tall and Vinner (1981) on concept image and concept definitions. Different parts of the concept image are activated depending on what the student perceives he needs to access at a given time; students encounter confusion when two conflicting parts of the concept image are activated *simultaneously*. This conflict was evident in our research when students struggled to construct tangent lines that touched the graph *only* at the point of tangency while still maintaining a correct slope. These two portions of their concept image disagreed, and the learner was required to make decisions about how to handle this conflict. According to Vinner (1982), activities that stimulate this type of cognitive interference are beneficial for the learner when handled properly.

Methods

The participants in this study were twelve first-semester calculus students at a large research university in the United States. Volunteers were sought to participate in the research project. Twelve students consented to the study, agreeing to participate in a 20-30 minute out-of-class interview, which was videotaped for further analysis. Six participants were taking calculus for the first time and six had previously taken calculus at the college level or in high school. Of the students who had previously taken calculus, not all of them were familiar with Newton's method. During the interviews, the students were asked about prerequisite knowledge related to Newton's method, such as how to construct a tangent line, identify *x*-intercepts, and interpret approximations. They were given a short reading describing the process of Newton's method and were asked to use Newton's method to graphically approximate the roots of a given function.

For this paper, we focus our analysis on students' graphical understanding of tangent line throughout the interview process, specifically attending to the different ways each student described a tangent line. The students' verbal and graphical responses from the interviews were organized into a chart format, which helped us to consolidate their responses and highlight emergent themes among the participants. We then compared students' graphical representations of tangent lines to key phrases used in their verbal descriptions and examined how these representations changed when using a tangent line within the context of Newton's method.

Results

Analysis showed six significant categories for students' concept image of a tangent line: *slope perspective*, *secant lines perspective*, *horizontal tangent perspective*, *one point perspective*, *multiple tangent lines perspective*, and *graph of* $y = \tan(x)$ *perspective*. Due to the limited space in this proposal, we chose to highlight three of these six categories. Several students showed evidence of using more than one perspective and were willing to change their concept image based on the task they were asked to perform.

Slope perspective

The *slope perspective* was the most prominent perspective occurring in the interviews, with nine of the twelve students being classified with this perspective. Students were classified as having the *slope perspective* if they sketched tangent lines by somehow considering the slope of the tangent line or the steepness of the graph at a given point. Interestingly, all of the students who consistently sketched correct tangent lines used the *slope perspective* at some point during the interview. Some of these students sketched a correct tangent line at a given point; others verbalized that they were considering slope, yet sketched a line that did not have the correct slope. Excerpts from interviews with three of these students are located in Table 1, below.

Table 1: Transcript Excerpts from Students Classified with Slope Perspective

Aaron	I'm just kinda guessing a little bit and putting it not so steep. … I'm not sure if it's actually supposed to be like down more, or actually sloping more.
Lewis (see Figure 1a)	[*looks at his first "tangent" line*] It should be probably like this [*draws a correct tangent line*] because the slope is going, it's, the slope is increasing so it actually should look like this.
Qadan (see Figure 1b)	Um, ok, the slope of the tangent line is the same as the slope of the point where it where it intersects.

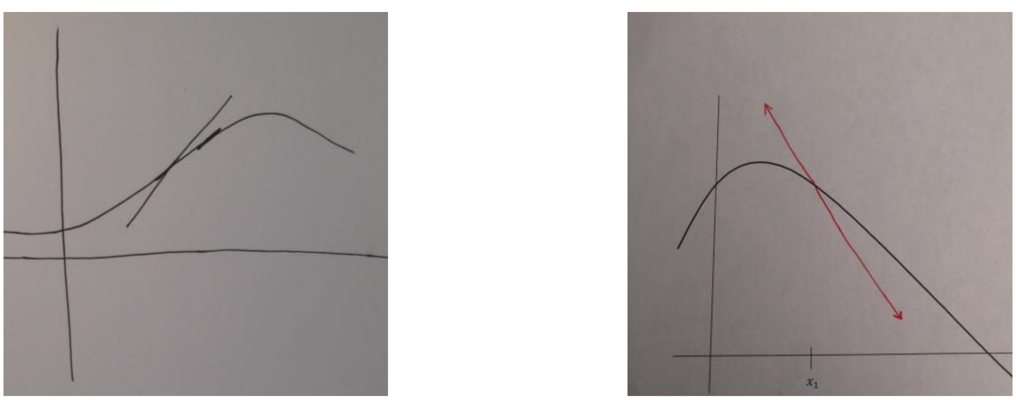


Figure 1a, 1b. Lewis' tangent line based on steepness of graph (left) and Qadan's tangent line based on the "slope of a point" (right).

Students had most recently been using tangent lines to sketch the graph of $f'(x)$ based on the graph of $f(x)$. For these tasks, students needed to recognize when the tangent lines had positive, negative, or zero slopes and whether the slopes were increasing or decreasing on an interval; they did not need to consider the exact placement of the tangent line. The interview forced students to use their existing understanding of tangent lines to draw a line tangent to a curve at a specific point. Some students were able to do this accurately, while others were unable to correctly accommodate their existing understanding of tangent lines with the act of drawing one.

Horizontal tangent perspective

The *horizontal tangent perspective* is a classification for students who, when asked to explain what a tangent line is, initially responded by describing or drawing a horizontal tangent line at a local maximum or minimum. Seven of the twelve participants were classified as fitting the *horizontal tangent perspective*. When analyzing this data, we hypothesized that the students responded with horizontal tangent lines because it was easiest for them to draw or explain in words. Upon further inspection, we saw students' verbal and graphical descriptions revealed that the horizontal tangent line played a larger role in how students were thinking about tangent lines. Three of the seven students referred to the horizontal tangent in their verbal definition of a tangent line. Example statements from their responses are given in Table 2.

Table 2: Excerpts from Students Using a Horizontal Tangent Line to Define Tangent Lines

Andrea	Isn't the tangent line whenever you're finding the derivative and that is the line that is horizontal to finding zero?
Bethany	When I think of, when I think of tangent line, I think of, uh, basically of it like being horizontal and that's where it's equal to the derivative.
Ted	Like, the *f* prime at *x* where the uh, where it equals zero. It's like this [begins drawing a horizontal tangent].

Andrea constructed only horizontal tangents with the exception of one incorrect attempt to mimic a non-horizontal tangent line drawn by the interviewer. Even when she was given a point on the graph whose actual tangent line would have a negative slope, she still drew a horizontal line through that point (See Figure 2). Her concept image of a horizontal tangent dominates her mental structure of tangent lines. She is situated in the assimilation process without accommodating her existing structure of a tangent line to include non-horizontal tangent lines.

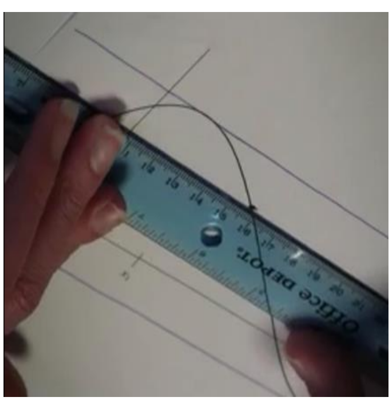


Figure 2. Andrea's horizontal "tangent:

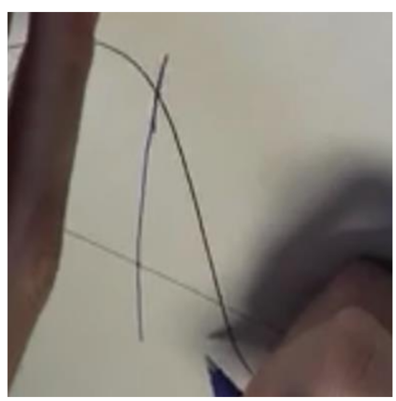


Figure 3. Aaron's result after "tipping over" his horizontal line

For Aaron, the horizontal tangent line was a starting point from which he constructed all other tangent lines (see Table 3 below and Figure 3 above). Aaron's graphical and verbal descriptions indicate that the horizontal tangent line influenced his understanding of tangent lines. Figure 3 shows the tangent lines Aaron drew from "tipping over" a horizontal tangent, an example of his use of assimilation by using his understanding of horizontal tangents to create a better understanding of other tangent lines. From his graph, we see that Aaron's adaptation of "tipping over" a horizontal tangent line is still not correct, but he was able to move slightly beyond Andrea, who believed that all tangent lines were horizontal lines drawn through a curve.

Table 3: Aaron's Explanation of How He Placed His Tangent Lines	
Aaron	'Cause if this is horizontal here, it's not going to get, you know, if it drops down, like straight then you know the tangent line would also be straight down if it was a really steep one like that. Then you know my tangent line would be steep…
Interviewer	So you're starting at the horizontal? And how do I get from the horizontal to this? [*points at non-horizontal tangent line*]
Aaron	Um yeah, I just kinda take the middle point of it and start moving it down and slowly I can imagine in my mind like tipping over a little bit.

One point perspective

Students classified as having the *one point perspective* were convinced that the tangent line must intersect the entire curve at exactly one point. A student with this perspective believes that tangent lines, when extended, may never intersect the curve again. The *one point perspective* is equivalent to Biza's (2010) "geometric global" category and is similar to Vivier's (2013) "global conception," where students attend to tangent lines globally and believe they cannot intersect the function at any point other than the point of tangency.

This misconception may be rooted in the terminology used by textbooks to introduce tangent lines. Kajander and Lovric (2009) note that several of the textbooks they studied used phrases like "touches at exactly one point" or similar wording (p. 175). In Biza's (2010) study, a large percentage of her students showed this sort of reasoning, while only two of our twelve students demonstrated this perspective. We believe the difference may be linked to Biza's students' familiarity with the concept of lines that are tangent to circles, which Biza stated was common in the students' curriculum. Tangents to circles were not a part of the curriculum in our calculus course, and when we asked our students about tangents to circles, most did not recall covering that material in high school. Another possible explanation for the infrequency by which our students used the *one point perspective* could be due to the instruction in the course. Since our instructors were aware of research on this common misconception, it is likely that they addressed this misconception in class. Nevertheless, we had two students who held to the *one point perspective.*

Mark demonstrated a strong understanding of a tangent line during his interview, but his *one point perspective* caused him to dismiss correct tangent lines as incorrect. He understood that the tangent line would be unique, and he did not try to alter the slope of the tangent line to make it fit his understanding. Rather, if the tangent line intersected the curve more than once, he believed there would not be a tangent line at that point. Thus, in Figure 4, Mark believed that the line drawn in the top graph would not be a tangent line, since "it can only intersect at like one point," but the line drawn in the bottom graph would be a tangent line. The *one point perspective* was a barrier to his understanding, but Mark was willing to accommodate his view of tangent lines after a quick explanation by the interviewer that dismissed this incorrect perspective.

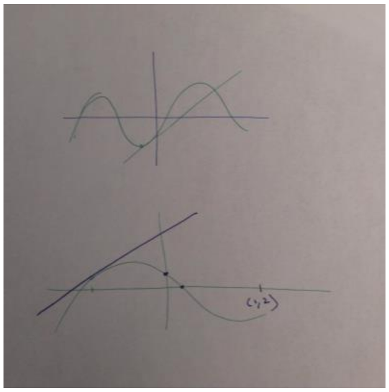


Figure 4. Mark's graphs using *one point perspective*.

Andrea was the second of our students who held to the *one point perspective*. She believed that if the tangent line hit the curve at another point, then the line would actually be a secant line. Notice in the excerpt below, Andrea was also using the *horizontal tangent line perspective*, and started by drawing a horizontal "tangent" line through a point on the graph whose tangent line should actually have a negative slope (See Figure 2).

> *Andrea:* Yeah. I'm still thinking it would still be like drawn this way [*draws a horizontal line*] oh no there's another point there [*indicates a second intersection point*] so it does have to be this way [*draws what resembles a normal line*] because you would run into a point here [*indicates the second intersection of the horizontal line*] and then that would make that the secant line instead of the tangent line.

Discussion

Recall that six prominent categories emerged from the data, but we only reported on three of those categories, due to the limited space in this proposal. Some of these categories were discussed in other research literature, but other categories that we found do not appear in the literature. A much larger study will need to be done to determine if these six categories are the only categories that arise from student thinking (we hypothesize that there are others that we did not see). At this stage, our research brings together several articles in the field as well as contributes to the research community's knowledge of how students think about tangent lines. Throughout the interviews, the students were usually willing to modify their existing concept image to accommodate new information. They did not have a fully developed tangent line structure, and therefore, learning new concepts such as Newton's method posed difficulties for them. Students were simultaneously working to take in new information (Newton's method) while still working to develop their concept image of tangent line. This process involves both assimilation and accommodation of material as new information was being processed and old information (tangent lines) was modified and adjusted to form a structure that makes sense for both conceptual domains.

When initially asked, "What is a tangent line?" most students used a piece of paper to sketch graphical examples of what they believed were tangent lines. One student, Bethany, seemed willing to change her understanding of tangent lines to match the information she was given about Newton's method, which suggests that the concept of a tangent line was far from solidified for her. When she was told to completely ignore the rest of the problem and just sketch a tangent line at a given point, she did it perfectly. However, as soon as she returned to the context of the problem, she began sketching incorrect tangent lines again. Likewise, after making several initial mistakes, Aaron drew a correct tangent line when he was asked to sketch tangent lines at specific points on a graph. Later, however, when he was sketching tangent lines to approximate the root of a function, Aaron drew tangent lines that ran through the curve (See Figure 3) rather than alongside the curve. Zandieh (2000) explains this phenomenon, noting that “students do not automatically connect an understanding of a process in one context with the same process in another context” (p. 125).

We suggest that instructors need to be aware that students hold multiple misconceptions about tangent lines and that the concept of a tangent line to a curve is perhaps not as simple for students as one might think. Phrases such as “touches the curve at only one point” are not sufficient explanation and can sometimes cause more confusion for students. It is worth the time to have students sketch several examples of tangent lines on multiple functions. Instructors should also note that, contrary to what might be expected, many students do not enter calculus having a strong concept image of a tangent line to a circle or any other graph, so these students are often constructing the concept image of a tangent line for the first time in calculus.

It is interesting to note that nine of the twelve students in the study were classified as using the *slope perspective*, making this the most common perspective that appeared during the interviews. This is consistent with Zandieh’s work (Zandieh, 2000), which examined student understanding of derivative in several contexts including slope, rate of change, physical contexts, and the difference quotient. She found that the graphical representation of slope was “the most frequently mentioned interpretation with six [of the nine] students mentioning it most often” (p. 119). Zandieh also noted that “beginning students’ preferences are not uniform but that they become more similar as students’ knowledge of the concept increases” (p. 119). Our research shows that this holds true for students’ concept images of tangent lines, as well. Interestingly, all of the students who consistently sketched correct tangent lines used the *slope perspective* at some point during the interview.

Implications for Teaching

As a starting point, instructors should be aware that tangent lines can be conceptually difficult for students. Phrases such as “touches the curve at only one point” are not sufficient explanation and can cause confusion for students. Instructors should also note that, contrary to what might be expected, many students do not enter calculus having a strong concept image of a tangent line to a circle or any other graph, so these students are often constructing the concept image of a tangent line for the first time in calculus. Further implications for teaching are expanded upon in another article that is under review at this time (LaRue et al).

Funding

Partial support for this work was provided by the National Science Foundation's ADVANCE IT Program under Award HRD-1007978. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Biza, I. (2008). *Models of students' conceptions about the tangent line.* Paper presented at the 11th International Conference on Mathematics Education, Monterrey,Mexico.

Biza, I., & Zachariades, T. (2010). First year mathematics undergraduates’ settled images of tangent line. *Journal of Mathematical Behavior*, *29*(4), 218-229.

Kajander, A., & Lovric, M. (2009). Mathematics textbooks and their potential role in supporting misconceptions. *International Journal of Mathematical Education in Science & Technology, 40*(2), 173-181.

LaRue, R., Vincent, B., Sealey, V., & Engelke, N. (under review). Calculus Students' Early Concept Images of Tangent Lines. *International Journal of Mathematical Education in Science and Technology*.

Páez Murillo, R. E., & Vivier, L. (2013). Teachers' conceptions of tangent line. *Journal of Mathematical Behavior, 32*(2), 209-229.

Piaget, J. (1970). *Structuralism*. New York: Basic Books, Inc.

Piaget, J. (1975). *The Equilibration of Cognitive Structures*. Chicago: The University of Chicago Press.

Tall, D., & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with Particular Reference to Limits and Continuity. *Educational Studies in Mathematics, 12*, 151-169.

Vinner, S. (1982). *Conflicts Between Definitions and Intuitions--the Case of the Tangent.* Paper presented at the Sixth International Conference for the Psychology of Mathematical Education: 6th Conference of the International Group for the Psychology of Mathematical Education.

Vivier, L. (2013). Without derivatives or limits: from visual and geometrical points of view to algebraic methods for identifying tangent lines. *International Journal of Mathematical Education in Science & Technology, 44*(5), 711-717.

Zandieh, M. (2000). A Theoretical Framework for Analyzing Student Understanding of the Concept of Derivative. In E. Dubinsky, Schoenfeld, A., Kaput, J. (Ed.), *Research in Collegiate Mathematics Education, IV* (Vol. 8, pp. 103-127). Providence, RI: American Mathematical Society.

COGNITIVE PROCESSES AND KNOWLEDGE IN ACTIVITIES IN COMMUNITY COLLEGE TRIGONOMETRY LESSONS

Linda Leckrone
University of Michigan
lleckron@umich.edu

Vilma Mesa
University of Michigan
vmesa@umich.edu

Over 50,000 students take trigonometry at two-year colleges in the U.S., yet little is known about their instruction. We report an analysis of activities in trigonometry classes taught at a community college attending to two dimensions, the type of knowledge used (Factual, Procedural, Conceptual, and Metacognitive) and the cognitive processes (Remember, Understand, Apply, Analyze, Evaluate, Create) intended in the activity as enacted by teachers in their lessons. Most of the 163 activities were classified as applying procedural knowledge; over one-fifth of the activities were coded as remembering factual knowledge or understanding conceptual knowledge. We discuss these findings in light of the community college setting and offer some questions for further research.

Keywords: Lecture, Trigonometry, Instructional Activities and Practices, Community College

Many post-secondary institutions offer trigonometry as part of a sequence of preparatory courses that lead to a calculus sequence. As of 2010 nearly 56,000 students take trigonometry at two-year colleges (Blair, Kirkman, & Maxwell, 2013), yet we know very little about how this topic is taught. As part of a larger study of mathematics instruction at community colleges, we collected a corpus of audio recordings of three consecutive trigonometry lessons. In this preliminary research report we discuss our analyses of the nature of the examples that teachers solved in the classroom. Specifically we sought to establish the type of knowledge and the type of cognitive processes that were predominant in these lessons. Our overarching research question is: "What are the types of knowledge used and the intended cognitive processes in activities enacted by these community college trigonometry teachers." We were curious to see whether, in this foundational course, students were being exposed to more than factual and procedural knowledge and whether the cognitive processes went beyond recalling and applying procedures. Most of the literature on mathematically demanding work suggests that students need to have challenging tasks in order to develop their understanding of mathematics. As part of this study we were interested in determining the extent to which trigonometry courses provided opportunities to learn rich mathematics for students seeking to major in a STEM degree.

Literature Review

Research on tasks used in mathematics classroom indicates that tasks that address novel mathematical questions are better at promoting student understanding than tasks that focus on routine or repetitive activities (Doyle, 1984, 1988). The way in which the tasks are enacted in the classroom matters. Students of teachers who systematically reduce the cognitive complexity of the tasks by asking simpler more routine questions perform worse on standardized tests than students of teachers who systematically maintain or increase the cognitive complexity of the tasks they work on (Silver & Stein, 1996; Stein, Grover, & Henningsen, 1996). One difficulty we had in characterizing task complexity at community colleges was due to the limited availability of instruments to capture elements of instruction in a mostly lecture-based setting. Anderson and colleagues (2001) proposed a revision of Bloom's taxonomy (Bloom, 1994) that provides a framework for analyzing the type of knowledge that can be elicited in an activity, as well as the different cognitive processes that students might engage in when working on the activity. Whereas the different types of knowledge are complementary—that is, one needs all of them in order to ensure an adequate knowledge of a subject—the cognitive processes differ in terms of the demand they impose on students and the amount of resources required (see Table 1).

Table 1: Definitions of the categories of the cognitive complexity coding scheme (Anderson, et al., 2001).

Type of Knowledge	Cognitive Processes Dimension
Factual Knowledge—Basic elements students must know to be acquainted with a discipline or solve problems in it, including knowledge of terminology and of specific details. ***Conceptual Knowledge***—Interrelationships among the basic elements within a larger structure that enable them to function together. It involves knowledge of classifications and categories, of principles and generalizations, and of theories, models, and structures. ***Procedural Knowledge***—How to do something, method of inquiry, and criteria for using skills, algorithms, techniques, and methods. It includes knowledge of subject-specific skills and algorithms, of specific techniques and methods, and of criteria for determining when to use appropriate procedures. ***Metacognitive Knowledge***—Knowledge of cognition in general as well as awareness of one's own cognition. It includes strategic knowledge, knowledge about cognitive tasks (including appropriate contextual and conditional knowledge), and self-knowledge.	***Remember***: Retrieve relevant knowledge from long-term memory, including recognizing and recalling. ***Understand***: Construct meaning from instructional messages, including oral, written, and graphic communication. It involves interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining. ***Apply***: Use a procedure in a given situation. It involves executing and implementing. ***Analyze***: Break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose. It involves differentiating, organizing, and attributing. ***Evaluate***: Make judgments based on criteria and standards. It involves checking and critiquing. ***Create***: Put elements together to form a coherent or functional whole and reorganize elements into a new pattern or structure. It involves hypothesizing, designing, and producing.

Methods

Participants

We analyzed 21 trigonometry lessons taught by five instructors of varying experience (See Table 2). The data was collected during Fall 2008 and Winter 2009 around the sixth week in the term, when most norms of classroom instruction had been established. Each instructor was observed for three consecutive lessons of each section taught (one of the instructors was teaching three sections). The lessons were audio-recorded and extensive field notes were taken documenting how the students were seated, who spoke in the classroom, and what was written on the board. The audio of the lessons was transcribed and the field notes added to the transcript to create a full description of the lesson.

Table 2: Instructor Characteristics

Instructor	Academic Background	Years of college teaching experience	Status
Ed	Mathematics, BS, MS	3	Part-time
Elizabeth	Mathematics, BS, MA	7	Full -time
Elliot	Economics, BS	6	Part-time
Emmett[a]	Physics, PhD	16	Full-time
Evan	Physics, BS; Mathematics, BS; Mathematics, MA	8	Part-time

a. Emmett taught three sections of the course.

Analysis

Transcripts and field notes were first reviewed to identify the examples teachers were using during instruction. We called these examples 'activities' or 'tasks' and used them as our unit of analysis. The first author made the identification, which was then verified and discussed with the second author. In general activities included discussion of examples in which there was an active exchange between students and teachers around the mathematics. We excluded activities that were not done with the whole class (e.g., discussion of problems with individual students at the beginning or end of the lesson). Activities were then coded along the two dimensions of the taxonomy. In coding the activities we considered the extent to which material was new to the students, the nature of students' questions or responses to teachers' questions, and teachers' language. It was frequent for teachers to emphasize different aspects in one activity; hence each activity could be assigned more than one of the codes in each category. The first author coded all the activities. During regular meetings with the second author all coding challenges were resolved and used to clarify the way in which the activities were coded. Below is an example of an activity that was coded as Factual, Remember. In this excerpt the teacher is emphasizing basic trigonometric angles that he expects students to know:

> *Teacher*: (writes on board) Sine of *x* equals square root of 3 over 2 and you guys know, I think by now, the sine of 30, 45 and 60, which one of those equals square root of 3 over 2?
>
> *Student*: 60.
>
> *Teacher*: Sine of 60 degrees. So *x* has to be 60 degrees, if we're talking about degrees. If we're talking about radians?
>
> *Student*: Pi over 3. (Elliot, 322-328)

In contrast, the following activity was coded as Procedural, Apply. Here, the teacher solved a problem from the book and summarized what had been done, emphasizing that using identities will "always work" in solving this type of problem:

> *Teacher*: … But this, the method of using trig identities does not rely on special triangles. … I repeat, this always works. The triangle only works if you recognize the triangle or is it a special type of triangle like 3-4-5. 12-5-13, that's a special triangle too. 6-8-10 triangle. But if you have different numbers it won't work. However, the identities always work. (Emmet, 230-239)

Results and Discussion

Altogether in these trigonometry lessons there were 163 activities, which received 207 codes. After this analysis we found that tasks that emphasized application of procedural knowledge were the most frequently used (57%), followed by tasks that emphasized understanding or applying conceptual knowledge (17%) and applying or remembering factual knowledge (16%). There were a small percentage of tasks in other codes (such as understanding procedures, 4%). No activities were coded as Create with any type of knowledge (see Table 3).

One way to interpret these results may be in the context of teachers' concern for students' affect. Asking students to create new knowledge can be unsettling for students and it has been documented that in general, mathematics teachers in the community college setting tend to be concerned with the affective well-being of their students (Mesa, 2012). Thus teachers tend to offer activities that are within students' level of understanding and to avoid activities that require struggles that may undermine students' sense of mathematical capacity. Given that community colleges enroll a disproportionate number of students who need remediation (Lutzer, et al., 2007), this concern is easily understandable.

Table 3: Number and Types of Tasks per Instructor

				Emmett						
	Ed	Elizabeth	Elliot	1	2	3	Evan	Total	% of codes	% of tasks
Factual Remember	1	2	5	1		1	5	15	7%	9%
Factual Understand	1						1	2	1%	1%
Factual Apply	1	1	4	3	1		8	18	9%	11%
Conceptual Understand	2	1	0	5	5	10	1	24	12%	15%
Conceptual Apply	1	1	1		5	3	2	13	6%	8%
Procedural Remember			2	1				3	1%	2%
Procedural Understand			1	1		3	3	8	4%	5%
Procedural Apply	12	20	27	31	10	9	10	119	57%	73%
Procedural Analyze						1		1	0%	1%
Metacognitive Apply		1			1			2	1%	1%
Metacognitive Evaluate		1				1		2	1%	1%
Total codes	18	27	40	42	22	28	30	207		
Total tasks	18	21	33	34	21	21	15	163		

The large number of activities coded as applying procedures suggests to us that this category will need to be re-examined, as it does not discriminate the different ways in which these teachers are presenting these procedures. In addition, we did not look at the entire lecture. We chose to look at only at activities because they seem to offer students the most opportunity to interact with the material. Yet this choice does not tell the whole story of what is going on in the classroom because teachers may have emphasized other cognitive processes and dimensions of knowledge in other parts of their lectures. Another area of interest is the quality of problems in textbooks used to teach this content. We found that our teachers chose a majority (72%) of their activities from the course textbook; as there is evidence that the examples available in these textbooks tend to be of low cognitive complexity (Mesa, Suh, Blake, & Whittemore, 2012) part of these results might be due to the low complexity of the tasks in the textbooks.

Questions:

1) Analyzing primarily lecture-based instruction is challenging. We chose to analyze activities because this offered the most opportunity for student participation and learning. Clearly, students may participate (take notes) in other parts of a lecture or they may ignore an instructor doing examples at the board. Are there other ways to get at opportunities for student learning in a primarily lecture-based classroom?

2) The method that we chose to use for analysis was not specific to mathematics. Other methods of analysis considered (Stein, Grover, & Henningsen, 1996) were specific to mathematics, but tended to focus on student work. As community college instruction is primarily lecture-based, we did not feel this was an appropriate method. Are there other types of analysis or methods that could have been used instead?
3) These instructors overwhelmingly chose to present activities in ways that emphasized applying procedural knowledge. Trigonometry tends to be a foundational course for other courses. Is there a way in which this method of presenting activities might not be seen as a problem?

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing.* New York, NY: Longman.

Blair, R., Kirkman, E., & Maxwell, J. (2013). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2010 CBMS Survey*. Washington, DC: American Mathematical Society.

Bloom, B. (1994). Reflections on the development and use of the taxonomy. In L. W. Anderson & L. A. Sosniak (Eds.), *Bloom's taxonomy: A forty-year retrospective, Ninety-third yearbook of the National Society for the Study of Education* (pp. 1-27). Chicago: University of Chicago Press.

Doyle, W. (1984). Academic tasks in classrooms. *Curriculum Inquiry*, 14, 129-149.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. Educational Psychologist, 23, 167-180.

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 CBMS Survey*. Washington, DC: American Mathematical Society.

Mesa, V. (2012). Achievement goal orientation of community college mathematics students and the misalignment of instructors' perceptions. *Community College Review*, 40(1), 49-77.

Mesa, V., Suh, H., Blake, T., & Whittemore, T. (2012). Examples in college algebra textbooks: Opportunities for students' learning. *Problems, Resources, and Issues in Mathematics Undergraduate Studies (PRIMUS)*. doi: 10.1080/10511970.2012.667515

Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education*, 30, 476-521.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33,* 455-488.

CALCULUS INSTRUCTORS' RESOURCES, ORIENTATIONS AND GOALS IN TEACHING LOW ACHIEVING STUDENTS

Misun Lee
University of Oklahoma

Sepideh Stewart
University of Oklahoma

Teaching and learning calculus has been the subject of mathematics education research for many years. Although the body of research is mainly concerned with students' difficulties with calculus, in this study we will be focusing our attention on the professors and instructors of calculus. In this research we used Schoenfeld's framework to examine four instructors' resources, orientations and goals in teaching calculus to low achieving students. So far, the preliminary results of the interviews show that although the professors thought differently about many aspects regarding teaching calculus, they all claimed that the first step to succeed in calculus courses is being prepared and having the right background.

Keywords: Calculus, Resources, Orientations, Goals, Low Achieving Students

Introduction

Calculus has been acting as a critical filter for many careers and continues to play a major role in STEM subjects. For a few decades mathematics educators have been concerned with teaching and learning of calculus. Artigue (2000) listed and discussed many difficulties that students have with calculus and considered the historical development of the curriculum to suggest ways of improving the current teaching. Norman and Prichard (1994) were alarmed that if the reports regarding the learning of calculus coming from various institutions around the United States were true "this country is in an abysmal state" (p. 65). The authors used Krutetskii's (1976) idea of *flexibility*, *reversibility* and *generalization* together with research on cognitive obstacles as a framework to understand students' difficulties in calculus. They found that the particular cognitive obstacles were very much tied to the state of mathematics instruction and suggested a reform of the mathematics curricula particularly in calculus. Robert and Speer (2001) believed that students' difficulties with calculus was universal and divided the research available in calculus into three categories of a) theory-driven, b) practice driven and c) convergence of the two. They believed that "the field will make progress on effective teaching and learning if it deals meaningfully with theoretical and pragmatic issues simultaneously" (p. 297). More than a decade later, has the research in calculus made any progress? Recently a large-scale survey of Calculus I was performed by the Mathematical Association of America (MAA) (Bressoud, Carlson, Pearson & Rasmussen, 2012). Although the study has given some insights into why students decide to opt out of calculus, the authors are still working on the analysis of the data and are not ready to make definite conclusions as yet.

Although, there are some research available on calculus students' difficulties, research on mathematics professor's day to day activities is scarce (Speer, Smith and Horvath, 2010). In response to this need, Sofronas and DeFranco (2010) did an extensive research to explore the knowledge base for teaching (KBT) among seven college and university mathematics faculty teaching calculus at 4-year institutions in the Northeastern United States. The authors developed a KBT framework among mathematics faculty teaching calculus. One of their findings was that "in the absence of any formal knowledge of learning theory, participants developed implicit "self-created" theories of student learning which influenced their teaching practices" (p. 193). To take the research one step further, in this study we are applying Allen

Schoenfeld's (2010) framework to examine calculus instructors' resources, orientations and goals.

Theoretical framework

The theoretical aspects of this study are based on Schoenfeld's (2010) recourses, orientations and goals (ROGs). He claims that "if you know enough about a teacher's knowledge, goals and beliefs, you can explain every decision that he or she makes, in the midst of teaching"(2012, p. 343). By resources Schoenfeld focuses mainly on knowledge, which he defines "as the information that he or she has potentially available to bring to bear in order to solve problems, achieve goals, or perform other such tasks" (p. 25). Goals are defined simply as what the individual wants to achieve. The term orientations refer to a group of terms such as "dispositions, beliefs, values, tastes, and preferences" (p. 29).

Although, the theory was originally considered as applying to research on school teaching, (Aguirre & Speer, 2000; Thomas & Yoon, 2011; Törner, Rolke, Rösken, & Sririman, 2010), it clearly has applicability to research on university teaching (Hannah, Stewart & Thomas, 2011; 2013; Paterson, Thomas & Taylor, 2011).

Based on Schoenfeld's theory we have developed a framework (see Table 1) to examine instructors' ROGs while teaching calculus. Our research questions are: What are instructors' resources, orientations and goals in teaching calculus courses? How does knowing teachers' ROGs result in helping the low achieving students?

Method

The research described here is a case study and it is as part of the first author's PhD thesis conducted at a large research University in 2013. The four participants (F1, F2, F3 and F4) in this research were tenured professors and an undergraduate advisor who have taught calculus courses in this research university more than two semesters. In order to fully explore the instructors' thought processes, three types of data related to the instructor's ROGs were collected: classroom observation field notes, instructor interviews, and course curriculum and information. The semi-structured interviews were from 40 to 90 minutes long. They were audio recorded and later transcribed. Some of the interview questions were: What do you expect students to learn from your calculus course? What do you think about investing class time to review and repeat problems to help students who do not follow the lecture easily? What is the level of mathematical understanding of your class? How do you evaluate yourself as an instructor in terms of dealing with low achieving students? Have you ever used technology to teach a calculus course? If yes, what kinds of technology have you used? Do you think using technology is helpful in teaching calculus? Inductive analysis approach was applied to the transcriptions of the interviews. The key step to analyze qualitative data was the tasks of comparing, contrasting, aggregating and ordering. The data were coded and analyzed based on the proposed framework. So far some of the emerging themes are: instructors' knowledge and views toward students; their effective teaching methods and orientations toward the curriculum and calculus. In the next section we will illustrate a brief account on the data that has been analyzed so far.

Preliminary Results

For the purpose of this proposal we have organized our preliminary result section into the following four categories:

Calculus

All four participants in this study reported that calculus is a subject requiring strong prior knowledge to build on. Even though they had differences in other areas regarding teaching and learning calculus, all of the interviewees agreed that if the students do not have adequate prior knowledge, there is nothing the instructor can do for them. Therefore, being in a right class is crucial to students. One instructor informed that many students desire to be in a higher level course:

Table1. A Framework to Illustrate Calculus Instructor's ROGs.

Resources	Orientations	Goals
Intellectual resources · *Knowledge of the nature of understanding calculus* - Calculus requires conceptual understanding. - Instructor's awareness of issues with difficulty of calculus. · *Pedagogical content knowledge* - Instructor's knowledge regarding most useful forms of representation, an understanding of what makes the learning of specific topics easy or difficult. **Contextual resources** · *Curriculum and content knowledge.* - Instructor's awareness of calculus curriculum and school policies. **Knowledge of students** · *Knowledge of student's goals* - Why do students take a calculus course?	**Pedagogical orientations** · *Orientation about calculus lectures* - Calculus lectures may design in such a way to allow students to solve real life problems as well as gain understanding by discovery. · *Orientation about technology* - Critical reflection is necessary while using the technology. - The role of technology must be clear in order to assist students' learning. **Learning orientations** · *Orientations about learning calculus* - To learn the subject well, students must be attentive and have adequate exposure to variety of problems. - Calculus students must have basic trigonometry as pre-knowledge.	**Pedagogical goals** · *Goals of calculus lectures* - Calculus lectures have to be motivating and act as a foundation for advanced mathematics courses. · *Goals of using technology* - The use of technology has to be based on a well-thought out pedagogical theory. - Calculus lectures may offer students opportunities for discovery using technology. **Subject matter goals** · *Goals of calculus concepts* - The geometrical representation of concepts must be clearly emphasized in lectures. - The material must be presented in such way that students gain conceptual understanding as well as adequacy to perform routine procedures.

F1: ... depending on the level the students to begin with it might not bring them up to the level they need to be. But it might be a start. And here again, that's why placement is so important placing the student in the correct course when they come in.

Students

Regarding the question, why do students take calculus courses? All the interviewees agreed that since calculus is a requirement for many other areas of science, students become less motivated in learning the concepts. They were also aware of many difficulties that students were facing while studying calculus. However, the common belief was that, students can improve their mathematics abilities if they work hard, for example, by reading a textbook, doing questions over and over again, and using other resources to get help.

F3: If they have questions or if they have difficulties they have to come up and be encouraged to ask. Things are not going to come to them. They have to come.

Teaching Goals

All of the instructors showed that the success of the course depended on the learner's conceptual understanding. They believed people would forget formulas but they won't forget concepts if they understand them well.

F4: You should understand why it is true and even you can't prove it later you should at least know that once you saw and understand why this is true.

One instructor also emphasized that calculus concepts are related closely to each other and exemplifying this relationship to students is her main teaching goal.

Effective Teaching Methods

Each instructor stated his/her own philosophy on effective teaching methods. They believed that using resources outside of the classroom such as mathematics help center, teaching assistants, and forming study groups would be helpful for many calculus students. Although all of the instructors emphasized that students should be more actively engaged in their learning, the way they encouraged students to utilize available resources varied by the individual. One instructor said:

F2: I gave her few extra points ... mostly I wanted to encourage her to keep coming back so she was given some small incentives to come here.

The instructors also expressed some limitations on helping students effectively, for example, factors such as teaching large size classes.

Questions

1. We would like your opinion on the proposed framework, specifically to know whether we have covered all the main resources, orientations and goals of a professor while teaching a typical calculus course.
2. While analyzing the data, deciding between orientations and goals of the professors has been challenging. We would like to know if anyone has had similar experiences in analyzing such data.

References

Aguirre, J., & Speer, N. M. (2000). Examining the relationship between beliefs and goals in teacher practice. *Journal of Mathematical Behavior, 18*(3), 327–356.

Artigue, M. (2000). Teaching and learning calculus: What can be learned from education research and curricular changes in France? *CBMS Issues in Mathematics Education*, Vol 8, pp. 1-15.

Bressoud, D., Carlson, M., Pearson, J.M. and Rasmussen, C. (2012). Preliminary results of the study on characteristics of successful programs in college calculus, ICME12, Topic Study Group 13 8 July – 15 July, 2012, COEX, Seoul, Korea.

Hannah, J., Stewart, S., & Thomas, M. O. J. (2011). Analysing lecturer practice: the role of orientations and goals. *International Journal of Mathematical Education in Science and Technology*, *42*(7), 975-984.

Hannah, J., Stewart, S., & Thomas, M. O. J. (2013). Conflicting goals and decision making: the deliberations of a new lecturer, In Lindmeier, A. M. & Heinze, A. (Eds.). *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 2, pp. 425-432. Kiel, Germany: PME.

Krutetskii, V. A., (1976). The Psychology of Mathematical Abilities in Schoolchildren, University of Chicago Press, Chicago.

Norman, F. A. & Prichard, M. K. (1994). Cognitive obstacles to the learning of calculus: A Kruketskiian perspective, In J. Kaput and E. Dubinsky (Ed.). Research issues in undergraduate mathematics learning, *The mathematics association of America*, Notes number 33.

Paterson, J., Thomas, M. O. J., & Taylor, S. (2011). Decisions, decisions, decisions: What determines the path taken in lectures? *International Journal of Mathematical Education in Science and Technology*, *42*(7), 985-996.

Robert, A. and Speer, N. (2001). Research on the teaching and learning of calculus/elementary analysis, In D. Holton (Ed.), *The teaching and Learning of Mathematics at University Level:* An ICMI Study, 283-299.

Schoenfeld, A. H. (2010). *How we think. A theory of goal-oriented decision making and its educational applications*. Routledge: New York.

Speer, N. M., Smith, J. P & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, 29, 99–114.

Thomas, M. O. J., & Yoon, C. (2011). Resolving conflict between competing goals in mathematics teaching decisions. In B. Ubuz (Ed.), Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education (Vol. 4, pp. 241-248), Ankara, Turkey.

Törner, G., Rolke, K., Rösken, B., & Sririman. B. (2010). Understanding a teacher’s actions in the classroom by applying Schoenfeld’s theory teaching-in-context: Reflecting on goals and beliefs. B. Sriraman, & L. English (Eds.), *Theories of Mathematics Education, Advances in Mathematics Education* (pp. 401-420). Berlin: Springer-Verlag.

DIAGRAMS IN ADVANCED MATHEMATICS: AFFORDANCES AND LIMITATIONS

We report a case study aimed at researching the rationale of a university mathematics professor for using diagrams in his analysis lectures, what he hoped his students would learn from these diagrams, the ways students understand these diagrams, and what they learn from them. Preliminary analysis suggest that by focusing on specific properties of the diagrams presented in mathematics lectures, or by attributing little importance to them, students fail to fully understand what professors hoped they would learn from these diagrams.

Key words: Proof, Diagrams, Mathematics Lecture, Visualization.

Kristen Lew	Tim Fukawa-Connelly	Juan Pablo Mejia-Ramos	Keith Weber
Rutgers University	Drexel University	Rutgers University	Rutgers University

Introduction

The undergraduate mathematics lecture is often criticized for focusing too specifically on the formal aspects of mathematics. One common critique challenges the notion that rigorous mathematical proof, at least as it is traditionally presented in mathematics lectures, is the best means of communicating mathematical explanations and justifications to students (e.g., Hersh, 1993; Thurston, 1994). The rigor present in these proofs can prevent students from having an intuitive understanding of why theorems are true (Hersh, 1993) and discourage them from using informal ways of understanding mathematics to construct proofs (e.g., Fischbein, 1982).

Thus, several researchers have called for undergraduate mathematics teaching to focus on more informal representational systems, specifically on the employment of diagrams and other visualizations to represent abstract mathematical concepts and proofs (Zimmerman & Cunningham, 1991). However, as pointed out by Speer, Smith, and Horvath (2010), there is very little systematic empirical research studying the actual teaching of university mathematics. In particular, there is little empirical data on how diagrams are being used in mathematics lectures, for what specific purposes, and what students understand from the use of diagrams in these presentations.

In this study, we consider the professor and students of an introductory undergraduate real analysis course and the diagrams encountered in a lecture covering the Riemann integral. As our theoretical orientation, we recognize that no matter what meaning instructors attempt to convey, students construct their own mathematical understandings (Thompson, 2013). We investigate these constructed understandings and how they align with the instructor's intended understanding. Specifically, we intend to provide insights into the following research questions: How and why are diagrams used? What types of mathematical insights do professors intend to convey to their students through the use of diagrams in lectures? How do students understand these diagrams? Do students gain the insights the professor hoped they would gain? If students do not gain these insights, what factors are inhibiting them?

Methods

In this study we consider one section of an introductory real analysis course taught at a large research university in the northeast of the United States.

Participants. This section of the course was led by a full professor with over 50 years of teaching experience at the university. This professor, who we will henceforth call Dr. A, was chosen for this study due to his reputation among his colleagues of being a careful and thoughtful lecturer, his very high instructional ratings in post-course student surveys, and for being an avid self-proclaimed proponent of using diagrams in his mathematics instruction.

Five students (two female, three male) were recruited from this class to participate in the study (students were modestly compensated for their participation). These students were pursuing either a major or a minor in mathematics (varying from first-year to fourth-year into their program) and demonstrated a range of mathematical aptitude according to the professor's assessment. We will use the codes S1, S2, S3, S4, and S5 for the five students.

Procedure. With the permission of Dr. A, the first author of this paper observed and videotaped one of his lectures. The first author then interviewed Dr. A to discuss his general view on the use of diagrams in mathematics instruction, his rationale for choosing the specific proofs, examples and diagrams he presented in the observed lectures and what he hoped to convey to his students in that lecture. Dr. A was then asked to watch a video clip of his presentation and explain why he had included each of the diagrams used, what he hoped to convey to his students through their presentation, and how he hoped his students would interpret each of these diagrams.

Next the first author individually interviewed each of the five students. These interviews focused on two of the diagrams Dr A. used in his lecture: a diagram used in the presentation of the definition of Upper and Lower Sums of the Riemann integral (Appendix A) and a diagram used in the presentation of a proof that the function f(x)=x has an integral from 0 to 1, which is equal to 1/2 (Appendix B). As an assessment of their understanding of the first diagram, each student was first asked to draw the upper and lower sums on the diagram of a different function/partition (Appendix A). Each student then discussed the definitions of upper and lower sums and how they related to the diagram they created, and what would happen if the partition were further refined.

Next, the students were asked to discuss the proof that the integral of the function f(x)=x exists and is 1/2. They were asked what they remembered about this proof, what they felt they learned from its presentation, and what they believed the professor was trying to convey through it. Each student then watched the video of the professor's presentation of this proof and was asked how the proof related to the diagram, and (once again) what they believed Dr. A was trying to convey through this presentation. Finally, as a way of assessing their understanding of this proof, each student was asked to attempt to prove that the integral of the function $f(x)=x^2$ exists and is 1/3, a proof that Dr. A indicated in his interview that his diagram presentation should enable students to tackle. Students then discussed their proof attempt and their use (or lack of use) of a diagram. At the end of the interview, students had an opportunity to discuss their thoughts on the use of diagrams in mathematics in general.

All interviews have been transcribed and all transcriptions have been summarized to give a general description of the participants' responses. Preliminary analysis of these responses has been conducted using the constant-comparative method (Strauss & Corbin, 1990).

Preliminary Results

Upper and lower sums diagram. In his presentation of the definition of Upper and Lower Sums of the Riemann integral, Dr. A used a diagram as a tool to highlight a concept that is central in the understanding of the definition. More specifically, he used the diagram to show what the least upper bound and greatest lower bound of a function looks like within an arbitrary interval of a partition. Furthermore, in a clear attempt to convey an intuitive understanding of what the upper and lower sums are, Dr. A explained to the class that "the upper sum will be an approximation which has a larger area than the actual area, and the lower sum will give you approximately a smaller area". Dr. A then used an analytic approach to represent the sums themselves.

In our interview, Dr. A said that he had used this diagram in particular to help the students remember the definition. Dr. A said: "the symbols by themselves do not produce in the mind a structure which can be easily remembered, so that, what you want to do is associate concepts' symbols with geometrical symbols." Further when asked what he wished to convey to the students with this diagram, he replied "the idea that the representation we're seeking is an approximation of area."

Of the five students, S1, S2, and S3 were able to successfully draw the upper and lower sums on the graph in Appendix A, while S4 and S5 were unable to complete the task successfully, even after watching the video of Dr. A's presentation of the definition. Further consideration of S4's and S5's work revealed serious confusion on the part of the students regarding both the definition and intuition of the upper and lower sums: S4 thought the upper and lower sums were simply the sums of the least upper bounds and the greater lower bounds, while S5 viewed the upper and lower sums as the areas above and below the function, respectively (as a consequence, S5 believed that a refinement of the partition would not alter the upper and lower sums at all).

Furthermore, when discussing how their own diagrams of the upper and lower sums related to the definition, it became clear that even S1 and S2 had some misconceptions. For instance, S1, S2, and S4 believed that the difference between the upper and lower sum yields the integral of the function.

The integral of f(x)=x proof. In his presentation of a proof that the function f(x)=x has an integral from 0 to 1, which is equal to 1/2, Dr. A drew the function, chose an arbitrary interval in a partition of *n* intervals of length *1/n*, pointing out what the least upper bound and the greatest lower bound were, and where they occurred on that interval. The proof then progressed in an analytic manner by setting up the sums, using series to solve them, and applying the Archimedean property to prove that the integral exists. In the interview, Dr. A explained that he had used this example and diagram as a "simple illustration of the concepts" presented in that lecture. Further he wished to convey the notion that "the approximation through upper and lower sums can actually be calculated." Dr. A also said he hoped the students saw the relevance of the graph as it is related to the proof.

When asked what they felt the professor was trying to convey through the presentation of this proof, each of the students responded that this proof was used to convey a new technique of finding an integral or to use the various concepts covered in class to prove an integral exists. After watching the video, S1 and S5 discussed the techniques and structure of proving that an integral exists, while S2 and S4 expanded on this pointing out that there are two parts of the proof, proving the integral exists and proving that the integral is equal to a particular number. However, after watching the video, S3 seemed more confused about the proof than when she had discussed it previously.

When asked what they believed Dr. A was trying to convey to them using the diagram, the students generally gave a description that was similar to Dr. A’s intended understanding. S1, S4, and S5 responded that the diagram was used to help students visualize the actual upper and lower sums. S2 said the diagram was used in order to aid the construction of the proof in a rigorous way. However, S3 said she was not sure what Dr. A intended to convey.

Of the five students, only S1 and S2 were partially successful at completing the proof that the integral from 0 to 1 of the function $f(x)=x^2$ exists and is equal to 1/3. Of these two successful students, only S1 drew a diagram for his proof. S2 reported that he visualized the graph while attempting to construct his proof and drew it subsequently. S3 outlined the structure of the proof without being able to do the necessary computations and neither S4 nor S5 were able to get beyond an attempt to set up the upper and lower sums. S4 and S5, however, did attempt to draw diagrams.

Thus, while students could state the purpose of the diagram in the lecture when specifically asked about it, the diagram was not useful to all of the students in understanding the content or in their proof-writing. In particular, S2 did not realize the professor had used a diagram in his presentation, S4 and S5 did not have the diagram in their notes (and could not remember much about it), and S3 was very confused by the diagram and did not understand its application.

Questions for the audience

What can we do to further analyze our data? Why do you think the use of diagrams was unsuccessful for students in this instance? Is there something that the professor could have done differently? What additional data might we want to collect from the students to help us decide whether they understand a diagram that an instructor presents in support of a proof?

References

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3* (2), 9-24.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24(4), 389-399.

Speer, Smith, and Horvath (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior, 29,* 99-114.

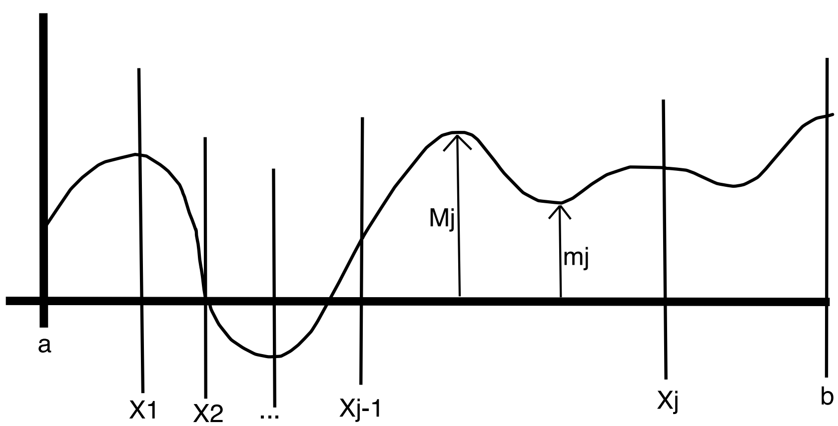
Strauss, A. & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* London: SAGE.

Thompson, P. W. (2013). Constructivism in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* [online]. Berlin: Springer. doi: 10.1007/SpringerReference_313210 2013-05-10 00:00:07 UTC.

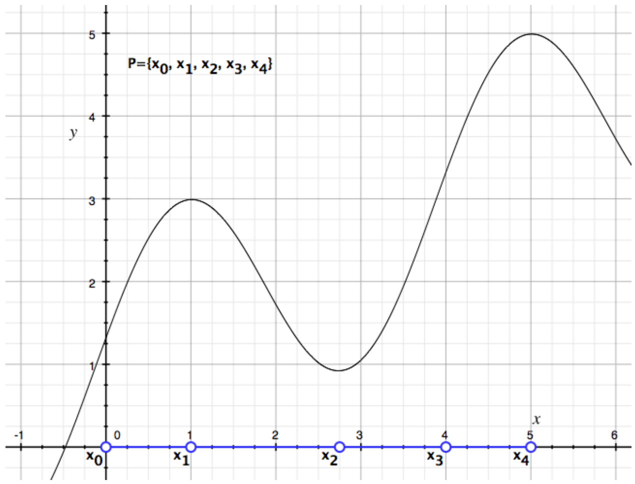
Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161–177.

Zimmermann, W., & Cunningham, S. (1991). *Visualization in teaching and learning mathematics.* Washington, D.C.: Mathematical Association of America.

Appendix A: Upper and lower sums diagram and task



a) Dr. A's Upper and lower sums diagram



b) Task for the upper and lower sums diagram. Students were asked to draw the upper and lower sums on the diagram.

Appendix B: Example proof diagram

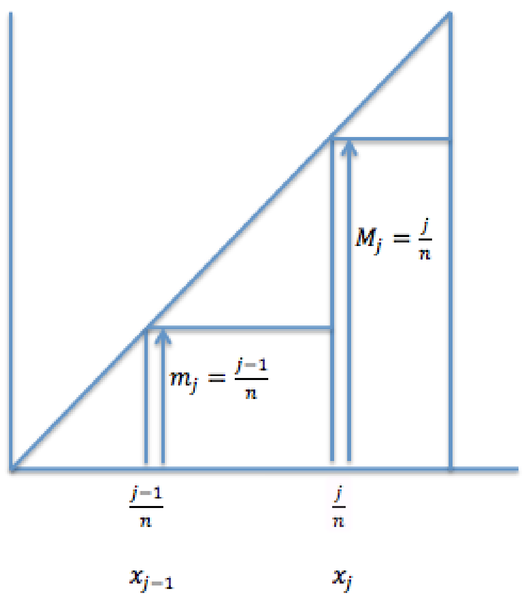


Diagram used by Dr. A in his presentation of the proof that the integral of the function f(x)=x exists and is 1/2.

WHY LECTURES IN ADVANCED MATHEMATICS OFTEN FAIL

Research on mathematicians' pedagogical practice in advanced mathematics is sparse. The current paper contributes to this literature by reporting a case study on a proof that a professor presented in a real analysis course. By interviewing the professor, we focus on his learning goals in this proof and the actions that he took to achieve these goals. By interviewing six students, we investigate how they perceived the proof and what they learned from it. Our analysis provides insight into why students did not learn what the professor desired from this lecture.

Key words: Lecture; Proof; Proof comprehension; Proof presentation

Many researchers in mathematics education claim that lectures in advanced mathematics do not enable students to build a robust understanding of mathematics. It is for this reason that Leron and Dubinsky (1995) asserted that there is a broad consensus amongst teachers and students that "the teaching of abstract algebra is a disaster and this remains true almost independently of the quality of the lectures" (p. 227). Indeed, "this is especially true for some excellent instructors" whose "lectures are truly masterpieces" (p. 227). Although these comments were specific to abstract algebra, Leron and Dubinsky's assertion that lectures often fail are applicable to lectures in any other subjects in advanced mathematics. However, Speer, Smith, and Horvath (2010) noted that these beliefs about lectures are usually based on popular opinion or personal opinion rather that systematic empirical study. There is little research on what goes on in lecture, what professors intentions are, and what students learn from lecture. The present study aims to fill this gap.

Related literature

Most studies on the teaching of advanced mathematics are composed of interviews with mathematicians about their teaching or self-reports in which mathematicians reflect on their own teaching (Speer, Smith, & Horvarth, 2010). The results of these studies are not presented here for the sake of brevity.

In terms of observations of teaching advanced mathematics, researchers primarily have used case studies. Author found that when one abstract algebra instructor presented proofs, she would model many of the mathematical behaviors associated with proof writing. She also consistently wrote out the logical details of the proof while only saying orally why some of these details needed to be justified (Author). In another study, Mills (2011) studied the different ways that a mathematician used examples to instantiate mathematical concepts. Author performed a semester-long case study on how one professor taught real analysis, regularly interviewing him about his teaching practices. He found the professor's practices were based on a coherent belief system and a good deal of thought, that like the professors studied by Author as well as Author, he would sometimes use informal representations of concepts such as examples and diagrams to help students understand the content.

Theoretical perspective

Purposes of proof

This study draws from two theoretical perspectives. First, we use de Villiers' (1990) purposes of proof to categorize what students might learn from a proof. According to de Villiers, proof can serve five purposes for mathematicians: to obtain conviction, to provide explanation, to promote discovery, to facilitate communication, and to systematize a theory. In the current study, systematization and communication did not play an important role in the proof that we studied; neither our research team nor the professor highlighted either role when discussing the proof. Hence, we will not consider these purposes of proof further in this proposal.

As deVilliers noted, an important function of proof is supplying conviction. If mathematicians are sure a proof is correct, they will feel compelled to accept the theorem being proven as true. However, studies with mathematicians reveal that conviction is not the primary reason that mathematicians read proofs (Author). Another purpose of proof is providing explanation. Hanna (1990) and Hersh (1993) argued that this, rather than conviction, should be the primary role of proof in the classroom. While there is not agreement on what constitutes an explanatory proof in mathematics education (Raman, 2003), here we follow Author in viewing a proof as explanatory for a student if that student can follow the proof in terms of informal representations (e.g., graphs) that are personally meaningful to him or her. A third purpose of proof is promoting discovery, where the prover or the reader can adopt the method of the proof to discover other theorems and proofs. According to mathematicians, this is the most important reason that they read proofs in the literature (Author) and exposing students to new methods of reasoning should be an important reason that students observe proofs in the classroom (Hanna & Barbreau, 2008).

Codes, competencies, and behaviors

In this paper, we follow the New Literacy Studies movement (Gee, 1990) and conceptualize the text of the proof as extending beyond symbols printed on a sheet of paper. Instead, we treat the totality of a lecture, including the oral words spoken by the professor, chalk inscriptions, and kinesthetic movements, as a single coherent piece of text. Treating the lecture as text allows us to use Weinberg and Weisner's (2011) framework for reading mathematics as a lens to interpret this data.

Like all genres of text, mathematics proofs contain *codes* that signify content. For instance, in the case of (written) proofs, Konior (1993) noted that an indented section of text indicates a sub-proof that can be read independently from the rest of the text. To comprehend text, students would need the *competencies*, or "the mathematical knowledge, skills, and understandings" (Weinberg & Weisner, 2011, p. 55), to understand these codes. Students would also need to engage in *behaviors,* or "sequences of actions (physical or mental) enacted by the implied reader" (p. 52), that would allow students to interpret this text in the way the author intends.

Research questions

This paper is based on a case study of a single proof in a lecture-based real analysis course. We study what content the professor intended to convey with this proof and what content the students gained from the proof. (We treat "content" broadly and refer to any insight or lesson that the professor attempted to convey or the student gained from reading the proof as content).Using the theoretical frames of deVilliers (1990) and Weinberg and Weisner (2011), we address the following research questions:

(1) What content did the professor hope to convey in the proof? Did it align with the goals of mathematics educators and include explanation and methods?

(2) How, if at all, did the professor hope to convey this content? Would it be recognizable to those enculturated into advanced mathematics?

(3) Did the students have the competencies to understand the lecture? That is, if shown a particular utterance or transcription from the professor, could the students decode what the professor was attempting to say?

(4) How did the students understand the proof that the professor presented? Did they have the understandings that the professor was intending to convey? If not, why not? Were their behaviors or competencies insufficient to comprehend the proof?

Methods

The proof

Our analysis focuses on an 11-minute proof about contractive sequences-- i.e., sequences with the property that $|x_n - x_{n-1}|<r^n$ for an r such that $0<r<1$. The theorem that was proven is that contractive sequences are convergent.

This proof was video-recorded during the eighth week of a 15-week semester. The videotape focuses exclusively on the actions of the professor, Dr. A (a pseudonym), and recorded what comments he said orally, what he wrote on the blackboard, and the gestures that he made.

The instructor's perception of the proof

The first author met individually with Dr. A for a 75-minute interview. During the interview, she first asked Dr. A why he chose to present this proof and what he was trying to convey to students through this proof. He was then asked to view the videotape of his lecture and stop the tape at every point at which he felt he was conveying content to the class and to describe what he was conveying. For each piece of content that Dr. A mentioned, we categorized the content as conviction, explanation, and method, as well, and how it was encoded.

Students' perception of the proof

Six students from the analysis course agreed to participate in video-taped interviews. The students were interviewed in pairs, on the rationale that this method might facilitate conversation between students and more authentic data (cf., Schoenfeld, 1985). Students were asked to bring their notes from the lecture to the interview, which were photocopied by the interviewer.

The interview format focused on students' understanding of the proof; four passes through the proof were made with each pair of students. First, students were shown a written copy of the proof (as Dr. A wrote on the blackboard) and asked what they thought Dr. A was trying to convey; the students were encouraged to consult their notes. Next, they were shown the proof in its entirety, asked to take notes and otherwise behave as they ordinarily would, and again asked what they thought Dr. A was trying to convey. Third, students were shown the eight videoclips that Dr. A highlighted as conveying information and were asked what they thought Dr. A was trying to convey (this pass directly investigated students' competencies at recognizing the codes that Dr. A used). Fourth, students were told that one purpose of the proof was to convey something that Dr. A had earlier identified as important and asked if they thought the proof accomplished this. We used a semi-open coding scheme, with the goals listed by Dr. A as *a priori* categories, to code students' responses.

Results

For the sake of brevity, we only highlight the main results; the presentation and full paper will discuss less central results as well as illustrate all results with transcripts. To answer each research question, the first was that the content that Dr. A aimed to convey primarily concerned method and conceptual explanation. Method content included providing a template for proving sequences converge; the importance of the triangle inequality; noting that one wants to show a sequence converges but does not have a candidate for limit, one can show the sequence is Cauchy; and expanding students' toolboxes to work with inequalities by keeping quantities small. For explanation, Dr. A sought to convey why the theorem was true with pictures representing convergent, contractive, and Cauchy sequences.

Second, Dr. A expressed all his content goals through oral comments. The only thing he inscribed on the blackboard was a fairly polished conventional proof. Some of his professed explanation goals, such as illustrating the theorem with pictures were accidentally omitted from the proof. For instance, when asked what he was trying to convey in the proof, he used the word “picture” 32 times in a six-minute period but when he saw the videotape of the proof, he said, “this a poor example. There are no pictures here! [laughs]”.

In the first two passes through the proofs with students, the students mentioned almost none of the content that Dr. A claimed he was trying to convey. Their performance improved somewhat in the third and fourth passes through the proof. For the third question, we note that students were able to discern some of Dr. A’s content, such as the importance of triangle inequalities and when one should prove sequences are Cauchy. However, other topics were more elusive. For instance, when students were asked about toolboxes of techniques to keep quantities small, students cited things such as using Mathematica and proof structures-- in other words, the toolbox to these students were any useful mathematical techniques. We noted that this tended to occur when the words that Dr. A signified conceptually rich mathematical structures and were repeated often by Dr. A. Students learned to repeat the words, but they did not signify the meaningful Students learned to repeat the words, but they did not use them in a way that signified the meaningful mathematics that Dr. A intended. For the fourth question, we noticed that one reason students did not cite content in the first two passes through the proof even though they had the competencies to decode it is that in the third pass five of the six students focused on the proof on the blackboard instead and did not record any of Dr. A’s oral comments. As oral comments were the way that Dr. A conveyed his content, this could account for why so much of the content was not recalled by the students.

Significance

As a case study, we cannot be sure what we observed will generalize to other professors or other proofs in real analysis with this professor. However, the themes that we observed offer guided hypotheses for why lectures are unsuccessful: professors sometimes unintentionally omit conceptual pictures from their proofs (see also Alcock, 2010; Author), students do not focus on ideas that are stated orally within proofs and most of the important content of the proof was presented orally by the professor, and some ideas, especially when short phrases signify rich mathematical structures, are not comprehensible to students in the way the teacher intends.

References

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in Collegiate Mathematics Education VII* (pp. 63-92). Providence, RI: American Mathematical Society.

de Villiers, M.D. (1990). The role and function of proof in mathematics. *Pythagoras, 24,* 17-24.

Gee, J. (1990). *Social linguistics and literacies: Ideologies in discourses.* Taylor & Francis: London.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6-13.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM,* 40, 345-353.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, *24*(4), 389-399.

Konior, J. (1993). Research into the construction of mathematical texts. *Educational Studies in Mathematics*, 24, 251-256.

Leron, U., & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly*, *102*, 227-242.

Mills, M. (2011). Mathematicians’ pedagogical thoughts and practices in proof presentation. In *Proceedings of the 14th Conference for Research in Undergraduate Mathematics*

Education. Edited by: Brown, S., Larsen, S., Marrongelle, K. and Oehrtman, M. Vol. 2, 283-297. Portland, OR.

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics*, 52, 319-325.

Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando: Academic Press.

Speer, N., Smith, J., Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior, 29*, 99-114.

Weinberg, A. & Wiesner, E. (2011). Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics*, 76, 49-63.

90

THE VALUE OF SYSTEMATIC LISTING IN CORRECTLY SOLVING COUNTING PROBLEMS

Elise Lockwood
Oregon State University

Bryan Gibson
University of Wisconsin-Madison

Abstract: Although counting problems are easy to state and provide rich, accessible problem solving situations, there is much evidence that students struggle with solving counting problems correctly. With combinatorics (and the study of counting problems) becoming increasingly prevalent in K-12 and undergraduate curricula, there is a need for researchers to identify potentially effective instructional interventions that might give students greater success as they solve counting problems. In this study, we tested one such intervention – having students engage in systematic listing of what they were trying to count. We found that even creating partial lists of the set of outcomes was a significant factor in students' success on problems. Our findings suggest that more needs to be done to refine instructional interventions that will facilitate listing. We discuss these findings, suggest follow-up studies, and request feedback from the audience.

Key Words: Combinatorics, Systematic listing, Counting problems, Experimental design

Introduction and Motivation

Enumerative combinatorics, or the solving of counting problems, has practical applications in probability and computer science, and it also provides a rich context for mathematical problem solving. As such, the solving of counting problems has become increasingly prevalent in K-12 curricula and in undergraduate mathematics courses. However, students tend to struggle with solving counting problems correctly, and there is a need for investigations into effective ways to improve students' counting. In this preliminary report, we share findings from a study that examined the effects of having students engage in *systematic listing* – that is, to create an organized list (or even a partial one) of the outcomes they are trying to count. We seek to answer the following research question: Does engaging in systematic listing have a significant effect on students' solving counting problems correctly?

Literature Review and Theoretical Perspective

For the most part, research on students' work on counting problems suggests that students struggle mightily with solving counting problems. Godino, Batanero, and Roa (2005) note that in Roa's (2000) study, 118 undergraduate mathematics majors "generally found it difficult to solve the problems (each student only solved an average number of 6 [of 13] problems correctly)" (p. 4). Eizenberg and Zaslavsky also reported low success rates, with their students correctly giving only 43 out of 108 initial correct solutions. Other researchers report specific mathematical features of counting problems that are especially difficult, such as issues of order (Lockwood, 2013; CadwalladerOsker, et al., 2012; Mellinger, 2004) and overcounting (Lockwood, 2011b, 2012; Annin & Lai, 2010). There is also evidence that students rely on memorized surface features of problems such as key words (e.g., Lockwood, 2011a), and that they struggle to know how to identify what a problem is asking (e.g., Hadar & Hadass, 1981). Mathematicians also acknowledge that counting can be difficult, and it is a domain in which "there are few formulas and each problem seems to be different" (Martin, 2001, p. 1). Given the overwhelming struggles that students seem to face when solving counting problems, there is a need to identify potentially productive interventions that may help students be more successful.

Theoretically, our focus on systematic listing stems from the idea that students may benefit from grounding their counting activity in the concrete set of outcomes they are trying to count. We hypothesized, based on prior experience and existing literature, that having students list might prevent them from blindly applying formulas and committing common errors of overcounting. This attention to the set of outcomes is suggested implicitly by a number of researchers (e.g., English, 1991; Hadar & Hadass, 1981; Mamona-Downs & Downs, 2004) and has been explicitly advocated by Lockwood (2011a, 2012). The study draws upon Lockwood (2013) model of students' combinatorial thinking, which proposes three basic components of students' counting (*expressions/formulas, counting processes*, and *sets of outcomes*) and elaborates on the relationships between these components. This model suggests that students may benefit from drawing upon the set of outcomes they are trying to count. The idea of systematic listing, and the act of reflecting on how to create an organized list of outcomes that correctly answers a counting problem, lies in the relationship between counting processes and sets of outcomes. Additionally, in a recent plenary address, Weber (2013) advocated for an increase in quantitative studies to complement qualitative studies in mathematics education. This study represents an attempt to respond to this call by providing preliminary quantitative data, which could better formulate subsequent questions that might be investigated qualitatively.

Methods

Forty-two undergraduate students participated in an hour-long written assessment. These students were enrolled in an introductory psychology course at a large Midwestern university, and they received extra credit in their course for their participation. Demographic information revealed varying degrees of experience with counting problems and suggested that almost all of the students had seen counting problems before (most typically in high school but not formally in college). The written assessment consisted of 9-12 counting problems, which were chosen based on the extent to which they might facilitate listing. We wanted some tasks that would encourage listing and that could be listed completely. We wanted other tasks with numbers large enough so as not to be able to be listed completely, to see if even partial listing might help students organize their work and detect patterns. We administered the surveys to students in each of three different iterations (the groups sizes were 13, 19, and 10, respectively), and we made some slight adjustments between each iteration to improve some of the tasks that seemed ambiguous or too difficult for students. In acquiring the data, we used Livescribe pens, which have technology that allows for written responses to be recorded in real time. These devices enabled us to examine how students formed their lists without conducting videotaped interviews.

The students were randomly assigned to either a *listing* or a *non-listing* condition, and each assessment involved pre-intervention, intervention, and post-intervention tasks. The tasks were the same for each condition; the only difference was in the prompts given to students. For the non-listing condition, students were simply asked to solve problems and show their work. For the listing condition, the students were additionally prompted with the following during the intervention tasks: "In the following 3 problems, please make an attempt to create a list of what you are trying to count." As will be discussed below, this prompt alone was not effective in actually getting students to list, but further studies could address particular ways in which listing might be elicited more effectively. An example of each type of task is given in Table 1 below.

Table 1 – Examples of tasks

Pre-intervention task	There are 5 different Spanish books, 6 different French books, and 8 different Russian books. How many ways are there to pick a pair of books not both in the same language?
Intervention task	You want to give 3 identical lollipops to 6 children. How many ways could the lollipops be distributed if no child gets more than one lollipop?
Post-intervention task	How many arrangements of the word CATTLE have the two Ts appearing together either at the beginning or the end of the word?

The first author coded the responses according to correctness (correct or incorrect) and according to four categories of listing (no listing, articulation, partial listing, and complete listing). A code of *no listing* was given if there was no attempt at any kind of partial or complete list. Typically a student who did not list would write down a numerical value or would write down some kind of formula or expression. A code of *articulation* emerged during analysis, as we realized that some responses were more than just providing a formula, but they were not quite suggestive of even a partial list. This *articulation* code was given when a student wrote down at least one instance of they were trying to count (one outcome) but did not create any kind of list. A code of *partial listing* was given if there was some evidence in the written work that the student was trying to create a list or partial list of the outcomes, but they may have not written the entire list correctly or may have truncated their listing when they identified a pattern. A code of *complete listing* meant a student provided a complete, correct list of the outcomes. Problems were coded one at a time to maximize the consistency in coding per problem. Qualitative analysis included watching through the students' work, determining features of *productive* (any lists, partial or complete, which were generated on a problem that the student eventually solved correctly) versus *unproductive* lists (lists that were written on a problem that did not have a correct answer), and identifying noteworthy aspects of listing that arose in the interviews.

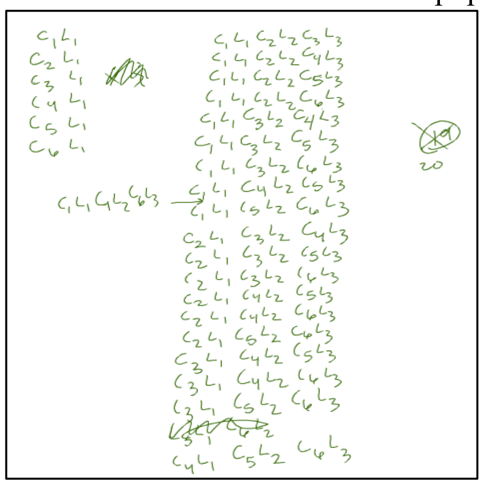
Results

Quantitative results. We combine all three iterations of the experiment for the following analysis, and only problems where the answer was clearly correct or incorrect, and where the listing behavior was clear were used (a total of 352 problems). On the whole, students struggled to solve these problems correctly, with only 24% (84/352) accuracy. We found that student performance on post-test questions, as measured by number of questions answered correctly, did not differ significantly between students told to list (intervention) and students who did not receive this instruction during the intervention phase. In other words, simply being instructed to list did not have a significant affect on future listing behavior. Additionally, if we look at the difference between the number of questions answered correctly in the pre-test versus in the post-test, we again do not see any significant difference between conditions.

However, if we look at listing behavior itself (and not whether students were instructed to list) we discover that listing had an overall positive effect on correctly solving a problem (here, we take listing as including a code of either partial or complete listing). If we count up the problems answered correctly by each student, the proportion of those where the student either partially or completely listed is significantly greater than those where they did not ($p < 0.02$). Simply using listing, then, seems to correlate positively with successfully answering a question, suggesting that listing may be a valuable counting activity in which students may engage. These results are discussed below.

Qualitative results. We also analyzed the students' written work captured by the Livescribe pens, and we distinguish between *productive* and *unproductive* lists, as defined above. Because of the nature of our data, we cannot make conclusive statements about whether or not a particular list actually caused a student to answer a problem correctly. However, we found the productive versus unproductive distinction to be helpful as we tried to determine potential aspects of listing that seemed particularly beneficial for students' counting. We identified three key features of productive lists: Useful notation and appropriate modeling of outcomes, an organized strategy, and evident structure. Student 121's work on the Lollipop's problem below (Figure 1) is an example of a productive list. She displays a notation that correctly models the outcomes as sequences of Cs and Ls (the Ls provide unnecessary information, but they do not distract her). The list is well organized, and the systematic way in which outcomes are listed helps her keep track of all of the outcomes and correctly determine the total number (to the point that she could identify an outcome that she had missed and correct an initial answer).

Figure 1 – Student 121's work on the Lollipop problem



Discussion

We note first of all that while there is a correlation between listing and correctly answering a problem, we do not claim causation. We acknowledge that it may be the case that stronger students may naturally list, and that is why we see the positive correlation. However, because of the overarching difficulties that students face with counting, we feel that the findings at least warrant more attention, and that the value of listing ought to be more carefully studied. Our findings suggest that we clearly were not careful or explicit enough in our instructions so as to make the intervention effective, and our encouragement to list was not consistently effective in getting students to list. However, given the potential benefit of listing, and given students' clearly documented struggles with counting problems, these initial quantitative findings suggest that there might be value in finding better ways to foster listing among students, and the preliminary nature of this report lends itself to follow up studies. Specifically, a major focus of future work is to investigate ways in which we can more effectively encourage students to list and better direct them in their listing activity. Such instructional interventions could be tested through interviews, through additional written assessments, and eventually in classroom settings.

Conclusions and Questions

Our aim in this study was to test whether or not listing might be a potentially helpful intervention. Our findings suggest that while our intervention was not entirely effective as worded, even partial systematic listing of outcomes did have a positive effect on students' correct solving of counting problems. This suggests that listing is a potentially promising aspect of work that could warrant further study, perhaps through subsequent in-depth qualitative studies. We pose the following questions to the audience:

- Given the nature of our current data, are there other questions that we should investigate?
- How else might we test the effectiveness of listing?
- Besides correctness, what are other potential factors would convince you that listing helps students count effectively?

References

Annin, S. A., & Lai, K. S. (2010). Common errors in counting problems. *Mathematics Teacher, 103*(6), 402-409.

CadwalladerOlsker, T., Engelke, N., Annin, S., & Henning, A. (2012). Does a statement of whether order matters in counting problems affect students' strategies? In the *Electronic Proceedings of the 15th Annual Meeting of the Research on Undergraduate Mathematics Education*. Portland, OR: Portland State University.

Eizenberg, M. M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning, 6*(1), 15-36.

Godino, J., Batanero, C., & Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60*, 3-36.

Hadar, N., & Hadass, R. (1981). The road to solve combinatorial problems is strewn with pitfalls. *Educational Studies in Mathematics, 12,* 435-443.

Halani, A. (2012). Students' ways of thinking about enumerative combinatorics solution sets: The odometer category. In the *Electronic Proceedings for the Fifteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education.* Portland, OR: Portland State University.

Lockwood, E. (2011a). *Student Approaches to Combinatorial Enumeration: The Role of Set-Oriented Thinking.* Unpublished doctoral dissertation, Portland State University.

Lockwood, E. (2011b). Student connections among counting problems: An exploration using actor-oriented transfer. *Educational Studies in Mathematics, 78*(3). 307-322. Doi: 10.1007/s10649-011-9320-7.

Lockwood, E. (2012). Counting using sets of outcomes. *Mathematics Teaching in the Middle School*, *18*(3). 132-135.

Lockwood, E. (2013). A model of students' combinatorial thinking. *Journal of Mathematical Behavior, 32,* 251-265. Doi: 10.1016/j.jmathb.2013.02.008.

Martin, G. E. (2001). *The Art of Enumerative Combinatorics*. New York: Springer.

Mellinger, K. E. (2004). Ordering elements and subsets: Examples for student understanding. *Mathematics and Computer Education, 38*(3), 333-337.

Roa, R. (2000). Razonamiento combinatorio en estudiantes con preparacion matematica avanzada. Unpublished Doctoral dissertation. Universidad de Granada.

Weber, K. (2013). Benefits and limitations of complementing qualitative research with quantitative studies: Lessons learned from four studies on proof reading. Presented at a plenary session of the *16th Conference on Research in Undergraduate Mathematics Education.* Denver, Colorado.

REINVENTING PERMUTATIONS AND COMBINATIONS

Elise Lockwood, Oregon State University
Craig Swinyard, University of Portland
John S. Caughman, IV, Portland State University

Abstract: Counting problems provide an accessible context for rich mathematical thinking, yet they can be surprisingly difficult for students. While some researchers have addressed these difficulties, more work is needed to uncover ways to help students count effectively. In an effort to foster conceptual understanding that is grounded in students' thinking, we had two undergraduate students engage in guided reinvention in a ten-session teaching experiment. In this experiment, the students successfully reinvented four basic counting formulas. In follow-up problems, combinations proved to be the most problematic for them, however, suggesting that the learning of combinations may require special attention. In this presentation, we describe the students' successful reinvention, and we discuss potential reasons for the students' issues with combinations. We additionally present potential implications and directions for further research.

Key Words: Combinatorics, Guided reinvention, Counting problems, Teaching experiment

Introduction and Motivation

Enumerative combinatorics has applications in probability and computer science, and its accessible yet challenging problems provide a rich context for mathematical reasoning. As a result, counting problems have gained traction in K-12 and undergraduate curricula in recent years, particularly in probability units and in discrete mathematics courses for undergraduates. Many researchers have described students' difficulties with counting problems (Batanero, Godino, & Navarro-Pelayo, 1997; Eizenberg & Zaslavsky, 2004) and have suggested a number of features that make the problems challenging (Hadar & Hadass, 1981; Martin, 2001). There remains a need for research that explicates how students can effectively comprehend basic counting principles.

The aim of our study was to gain insight into how students might come to reason coherently about four basic counting formulas: $n!$, n^r, $\frac{n!}{(n-r)!}$, and $\frac{n!}{(n-r)!\,r!}$. Textbooks typically present these early on, following each with numerous examples. Students are generally expected to apply them in various contexts throughout the remainder of the course. Research (Lockwood, 2013) indicates, however, that students frequently misapply these formulas, which suggests they may not understand how and why these expressions differ. We conjectured that engaging students in a *guided reinvention* of these basic counting principles could provide us an opportunity to understand how students make sense of them. Studies indicate (Swinyard, 2011; Swinyard & Larsen, 2012) that reinvention can be helpful for students to develop coherent reasoning, and also for researchers to gain insight into how students come to understand particular mathematical concepts. In this study, we engaged a pair of undergraduates in a ten-session teaching experiment, during which they solved basic counting problems and then subsequently generalized their mathematical activity by reinventing the four formulas above. In this paper, we report on the students' reinvention of the four formulas, addressing both of the following research questions:

1) *How might students reinvent these four basic counting formulas?*
2) *What cognitive issues might arise for students as they reinvent and use these formulas?*

Literature Review and Theoretical Perspective

The vast majority of research on the teaching and learning of combinatorics is relatively new. Thus far, researchers have clearly established that students have difficulty with even basic counting problems. For example, Eizenberg and Zaslavsky reported that in their study of undergraduates, "only 43 of the 108 initial solutions were correct" (2004, p. 31). Some researchers have identified common student errors, including over-counting and confusion about when order matters (Annin & Lai, 2010; Batanero, et al., 1997; Kavousian, 2006). Others have identified factors that might lead to such difficulties (Batanero, et al., 1997; Hadar & Hadass, 1981) and have identified potentially productive problem-solving and verification strategies (English, 1991, 1993; Eizenberg & Zaslavsky, 2004). More recently, Lockwood (2013) has proposed an initial model of students' combinatorial thinking and emphasized the importance of focusing on sets of outcomes when solving counting problems.

Research that emphasizes the student's perspective has proved useful in other mathematical content areas (Swinyard, 2011) and in combinatorics (e.g., Halani, 2012; Lockwood, 2013). Following Lobato (2003), Lockwood (2011) examined student-generated connections through an actor-oriented perspective, tracking what the students themselves saw as similar among counting situations, thus reorienting the perspective toward students' views of counting.

With a similar aim of emphasizing the student's perspective, we drew inspiration from the perspective of *developmental research* (Gravemeijer, 1998). Developmental research leverages students' informal knowledge and supports them in developing sophisticated, abstract knowledge while maintaining intellectual autonomy (p. 279). A heuristic commonly associated with developmental research is *guided reinvention*, "a process by which students formalize their informal understandings and intuitions" (Gravemeijer, Cobb, Bowers, and Whitenack, 2000, p. 237). The formalization process necessarily requires students to generalize their previous mathematical activity. In line with Freudenthal's recommendation (1973) to avoid an *antididactic inversion* (where symbolic formalism precedes reasoning), we aimed to create an environment that fosters initial exploration of counting problems that emphasizes sense-making over conventional symbolization. We conjectured that we might gain useful insight into how students come to understand basic counting principles if we engaged them in activities designed to foster their *reinvention* of the basic formulas.

Methods

The aim of this paired teaching experiment (Steffe & Thompson, 2000) was for the two students (Thomas and Robin) to reinvent four basic counting principles based on their engagement with a variety of counting problems. The motivation for the paired teaching experiment stemmed from the second author's experience with the reinvention of advanced calculus definitions (Swinyard, 2011; Swinyard & Larsen, 2012). The participants in this study were two above-average students who had recently completed an integral calculus course taught by the second author. They were chosen based on the following criteria: 1) no formal experience with combinatorics; 2) strong mathematical background and ability; and, 3) a propensity to engage actively with mathematics and articulate their reasoning, as observed by the second author during the integral calculus course.

The teaching experiment consisted of ten, 90-minute sessions and proceeded in two phases. During Phase 1 (Sessions 1-3) the students reasoned about and solved ten relatively elementary counting problems, thus providing them with a common experience from which they could later generalize. The solution set for most of the problems was small enough that the students could completely enumerate the solutions in a table or tree. During Phase 2 (Sessions 4-10), the students encountered more challenging tasks, both in terms of the size of the solution set, and in the sophistication needed to solve each task. The aim of Phase 2 was for the students to reinvent each of the four basic counting formulas. We intentionally increased the cardinality of the solution sets so that the students might be motivated to generalize their prior mathematical activity in a manner that supported them in efficiently solving similar counting problems.

The analysis of data occurred at multiple levels. As the teaching experiment proceeded, we conducted an *ongoing analysis* that included reviewing the videotape of each session and constructing a "content log." In creating these content logs, we paid particular attention to students' articulated thoughts that seemed to provide them with leverage, the voicing of concerns or perceived hurdles that needed to be overcome, and signs of/causes for progress. Our ongoing analysis informed our research team's decisions about tasks for subsequent sessions. We have also conducted a *retrospective analysis*, in which we review the entire corpus of data at a deeper level, so as to refine our descriptions of thematic elements present in the students' reasoning.

Results

In this section, we summarize the students' reinvention of the four fundamental counting formulas. At the outset of the teaching experiment, we gave the students ten counting problems, which they solved during the first three sessions. The students quickly adopted a consistent approach to solving each problem, in which they read the problem aloud, wrote down relevant information on the board, and then attempted to write a table of the outcomes they were trying to count. Remarkably (given documented low success rates, e.g., Eizenberg & Zaslavsky, 2004), despite having no previous formal experience with counting, the students correctly solved all ten problems on their initial attempt.

Initial problem solving. As an example of their typical method of operation, we briefly describe their work on one of the initial problems: *How many arrangements of the letters in the word CATTLE have the two T's appearing together either at the beginning or the end of the word?* To solve this problem, the students wrote out the word "CATTLE" on the board, and after some discussion decided that Thomas would list ways to arrange the letters with the Ts at the beginning, and Robin would list ways to arrange the letters with the Ts at the end. They developed systematic ways of listing possibilities, and each wrote all six ways of arranging the letters CALE with C as the first letter. Robin's summary of their solution process below captures the problem-solving approach they typically employed for the first ten problems.

> **Robin:** After doing the first few we realized the pattern. And so we saw that, if we start with the C, and do all the swapping, we get six combinations. And if we start with [A] we're going to get another six combinations…But if we start with L we have six more, if we start with E we have six more, totaling 24. And his 24 plus my 24 would make a total of 48.

Reinvention of formulas. After the first ten problems, the students had developed effective problem solving techniques and a good rapport with one another. However, we realized that their organized listing and pattern detection had not motivated a need for any more generalized methods or formulas. Since the aim of the teaching experiment was to gain insight into the students' reasoning as they reinvented four basic counting formulas, we recognized the need to provide tasks that would necessitate generalization for the students. Toward that end, we presented them with problems whose solution sets were too large to enumerate easily via listing. Table 1 shows the problems designed to facilitate the reinvention of each respective formula.

Through engaging with the problems, the students successfully reinvented each formula, using their own notation. For instance, the f in the last two formulas stood for fans at the basketball game.

Table 1

Problem	Textbook Formula	Students' Formula
In the downtown public library, there are 648 books in the children's section. In how many different ways can all of those children's books be arranged on the shelves of the library?	$n!$	$n!$
There are 40 houses in the neighborhood, and they each need to be painted this summer. There are 157 paint colors available. In how many different ways could all of the houses be painted?	n^r	a^b
There are 19,000 fans at a basketball game. During halftime, a first, second, and third place prize are going to be given out to three lucky fans. In how many ways can the prizes be given out? There are 19,000 fans at a basketball game. Throughout the game, fifty randomly chosen fans are going to be given fifty different prizes. How many possibilities are there for how the prizes can be distributed?	$\frac{n!}{(n-r)!}$	$\frac{n!}{(n-f)!}$
There are 19,000 fans at a basketball game. During halftime, three lucky fans get to participate in a free throw contest. How many possibilities are there for which fans can participate? There are 19,000 fans at a basketball game. After the game, fifty fans are going to be chosen randomly to meet the team. In how many ways can these fifty fans be chosen?	$\frac{n!}{(n-r)!r!}$	$\frac{n! \div (n-f)!}{f!}$

Exploring Choosing. Immediately following their reinvention of the formulas, the students successfully solved two problems involving each type of formula with varying contexts. For instance, they solved problems like *A fair coin is flipped 36 times. How many outcomes have a head on the fifth toss?* (n^r), and, *In a shipment of 1000 iPhones, 25 are defective. In how many ways can we select a set of 50 non-defective iPhones?* ($\frac{n!}{(n-r)!r!}$). Then, to our surprise, the students struggled significantly with another combination (or "choosing") problem: *Consider binary strings that are 256 digits long. How many 256-bit strings contain exactly 75 0s?* We designed this as a combination problem, because one can select 75 of the 256 spots to place the 0s and fill in the rest of the spots with 1s. The students spent over an hour on this problem, and for the first time in the teaching experiment they were unable to arrive at an answer.

We realized that they were not recognizing the problem as a combination problem, and, in an effort to guide them toward this realization (without explicitly telling them to select spots), we presented them with another problem that could be solved using the combination formula: *There are 40 houses in the neighborhood, and they each need to be painted this summer. There are 157 paint colors available. In how many different ways could all of the houses be painted, if exactly two of the houses must be red?* This problem can be solved by selecting which two of the 40 houses are red, and then painting each of the remaining houses any of the remaining colors. Thomas and Robin correctly noted that they could paint the non-red houses in 156^{38} ways. However, while they indicated a need to determine which two houses were red, they did not recognize that they could enumerate this task using their combination formula. They solved the problem using careful listing and patterns, recognizing what they called an "adding factorial" in the solution, as there are 40+39+…+3+2+1 total ways to select the houses. Thus, while they solved the problem correctly, they still did not recognize it as involving combinations.

At this point, then, we recognized that while the students had been overwhelmingly successful on many problems, there was something about the act of choosing that was problematic for them. In the next session, we found further evidence of their struggle with recognizing when and how to apply the proper formula in these combination problems. First, we gave them the problem: *There are 39 flamingos at the zoo. The zoo is going to exchange 6 of them with another zoo for 5 peacocks. How many possibilities are there for which flamingos are exchanged?* Interestingly, they solved this correctly, almost immediately using the combination formula. We then gave them another house problem, *If there are 40 houses on the block, in how many ways can exactly 3 houses be painted blue?*, and they did not make the connection and again solved the problem through listing outcomes and finding a pattern. The juxtaposition of these two problems is striking. From the expert's perspective, these problems are isomorphic, and yet the students clearly did not see them as such.

Discussion

A number of interesting discussion points emerged in the students' problem solving, most notably the two problems with which they struggled. We find it interesting that Thomas and Robin recognized when to apply the combination formula in some contexts but not in others. Without question, the students correctly applied the other three formulas with relative ease, but only made use of the combination formula in particular contexts. We have two conjectures that emerged from the teaching experiment as to why this might be the case.

First, it seems as though there were certain features of what was being chosen that affected the students' ability to recognize that the combination formula was applicable. At one point Robin stated that she saw the houses as being static – while she could envision people or flamingos actually moving and being chosen, she could not do the same with the houses. She additionally indicated that she interpreted the houses as being identical: "When I see people I think three different people, and when I see houses I think the same house painted blue three times." The different contexts appeared to elicit different features in the objects (distinguishability, dynamics, etc.) that influenced the students' ability to see the problems as involving the combination formula.

We also believe that there may be something fundamental about the abstract act of choosing that is related to reification. While the students at times had some notion that they were "picking" basketball fans to participate, they did not seem to view the combination formula as being applicable across a wide variety of contexts. While they could use the expression in a few directly worded problems, they did not seem to be able to abstractly map their own choosing of houses (or spots in which to place zeros) onto their understanding of the combination formula. We suspect there is something cognitively different between how the combination formula is used in these different contexts.

Another possible factor contributing to the students' struggle with certain combination problems is that their initial understandings of the formulas during their reinvention were not sufficiently conceptually grounded. Indeed, in our analysis of these problems we were struck by their reliance on patterns. We believe that they relied on what Harel (2001) calls result pattern generalization, and their formulas (and their explanations of those formulas) were almost entirely grounded in that pattern recognition. As an example, while we repeatedly asked them why they needed to divide by $f!$ in their formula, their arguments tended to be because "that was the pattern we saw" and not because they clearly understood the function the *f*! played. When pressed, they could articulate a vague conceptual explanation, but their default explanation was to argue based on the patterns. This was true even in the other formulas – indeed, rather than being grounded in a fundamental notion of multiplication, even their notions of permutations were surprisingly rooted only in the observation of a numerical pattern.

Conclusion and Future Directions

In this paper, we have described how students might reinvent basic counting formulas. In addition to providing insight into a potential trajectory through which students might reinvent counting formulas, an interesting aspect of students' thinking about combinations emerged. This work has raised questions that inspire subsequent studies, such as whether or not similar difficulties with combinations will arise with other students. If so, can we learn more about why choosing in some contexts is so difficult? Additionally, we are curious about how having students reinvent these four formulas might support students in handling problems with repetition, and in reinventing foundational counting principles (such as the addition and multiplication principles).

References

Annin, S. A., & Lai, K. S. (2010). Common errors in counting problems. *Mathematics Teacher, 103*(6), 402-409.

Batanero, C., Navarro-Pelayo, V., & Godino, J. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*, 181-199.

Eizenberg, M. M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning, 6*(1), 15-36.

English, L. D. (1991). Young children's combinatorics strategies. *Educational Studies in Mathematics, 22,* 451-47.

English, L. D. (1993). Children's strategies for solving two- and three-dimensional combinatorial problems. *Journal of Mathematical Behavior, 24*(3), 255-273.

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht, The Netherlands: Reidel.

Gravemeijer, K. (1998). Developmental research as a research method. In A. Sierpinska & J. Kilpatrick (Eds.), *Mathematics Education as a Research Domain: A Search for Identity* (pp. 277-296). Dordrecht, The Netherlands: Kluwer.

Gravemeijer, K., Cobb, P., Bowers, J., and Whitenack, J. (2000). Symbolizing, modeling and instructional design. In P. Cobb, E. Yackel and K. McClain (Eds.), *Symbolizing and Communicating in Mathematics Classrooms*, Erlbaum, Mahwah, NJ, pp. 225–273.

Hadar, N., & Hadass, R. (1981). The road to solve combinatorial problems is strewn with pitfalls. *Educational Studies in Mathematics, 12,* 435-443.

Halani, A. (2012). Students' ways of thinking about enumerative combinatorics solution sets: The odometer category. In the *Electronic Proceedings for the Fifteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education.* Portland, OR: Portland State University.

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. R. Campbell & R. Zaskis (Eds.), *Learning and teaching* *number theory: Research in cognition and instruction* (pp. 185-212). Journal of Mathematical Behavior Monograph. Westport, CT: Ablex Publishing.

Kavousian, S. (2008) *Enquiries into undergraduate students' understanding of combinatorial structures.* Unpublished doctoral dissertation, Simon Fraser University – Vancouver, BC.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17-20.

Lockwood, E. (2011). Student connections among counting problems: An exploration using actor-oriented transfer. *Educational Studies in Mathematics, 78*(3). 307-322. Doi: 10.1007/s10649-011-9320-7.

Lockwood, E. (2013). A model of students' combinatorial thinking. *Journal of Mathematical Behavior, 32,* 251-265. Doi: 10.1016/j.jmathb.2013.02.008.

Martin, G. E. (2001). *The Art of Enumerative Combinatorics*. New York: Springer.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Swinyard, C. (2011). Reinventing the formal definition of limit: The case of Amy and Mike. *Journal of Mathematical Behavior, 30,* 93-114.

Swinyard, C., & Larsen, S. (2012). Coming to understand the formal definition of limit: Insights gained from engaging students in reinvention. *Journal for Research in Mathematics Education, 43*(4), 465-494.

Noticing the Math in Issues of Social Justice

Ami Mamolo
University of Ontario Institute of Technology

This preliminary report examines pre-service secondary mathematics teachers' engagement with problems which contextualized mathematics in issues of social justice. A framework for Teaching Mathematics for Social Justice was employed and participant responses were analysed with respect to what mathematics they noticed and attended to in and after the problem solving. Results suggest participants had difficulty "seeing" the math in non-math contexts, and that their ability to notice the embedded mathematics was influenced by the specific social context as well as their orientation towards mathematics (both in general and regarding specific content). Implications for research and teacher education are described.

Key words: pre-service teacher education; social justice; context problems

Consider the following contexts:

- Delivering fresh foods to remote and/or inaccessible communities
- Informing policies on fair trade of (e.g.) chocolate products
- Investigating causes and consequences of the Savar building collapse

If you were asked to think about these issues of social justice from a mathematical point of view, what would come to mind? Would you notice specific mathematics concepts, such as optimization or cost analysis, or broader mathematical reasoning, such as spatial or numerical sense? What experiences would help with noticing, and mobilizing, the mathematics embedded in these issues?

These questions highlight the motivation behind the research presented in this preliminary report. Specifically, this research addresses the question: *What do pre-service teachers notice and attend to when exploring social justice context problems – that is, when they are asked to engage mathematically with issues of social justice?*

Mathematics lessons that teach to issues of social justice have been used recently to make math meaningful in multicultural classrooms (e.g. Bateiha, 2010), to improve student understanding of world issues (e.g. Bartell, 2011/2013), and to strive for equity in the classroom (e.g. Stinson, 2013). According to Gutstein's (2006) model for Teaching Mathematics for Social Justice (TMSJ), a balance must exist between Social Justice Pedagogical Goals (SJPG) and Mathematics Pedagogical Goals (MPG), but this balance can be difficult to achieve (Bartell, 2011/2013). Social justice context problems offer an avenue through which to achieve this balance, however as Beswick (2011) observes "enthusiasm for context problems appears to be in advance of the evidence for their effectiveness" in learning mathematics (p.387). This research is part of a broader study which explores how the design of social justice context problems, and how engagement with such problems, may influence understanding of mathematics and mathematics teaching.

Background and Theoretical Framework

The connection between mathematics learning and addressing important and relevant world issues related to social justice has begun to flourish in education research. Attention has focused on the possibilities and realities of learning mathematics for social justice, and research has ranged from developing pedagogy for multicultural settings (e.g. Bateiha, 2010) and investigating teacher challenges when negotiating pedagogical goals (e.g. Bartell, 2011/2013; Esmonde et al., 2013), to exploring the impact on student identity (e.g. Gutierrez, 2013) and their understanding of critical world issues (e.g. Gutstein, 2003, 2006). In contrast, the research presented in this proposal attends to learning mathematics *through* social justice issues – a dimension that has not yet been explored empirically in mathematics education research. For the purposes of this paper, the distinction between *for* and *through* is made to emphasize the scope of the research – it is on individuals' understanding of mathematics as evoked through issues of social justice.

The theoretical framework that underpinned the methodology and data analysis was that of Gutstein's TMSJ model. Its two components – Social Justice Pedagogical Goals (SJPG) and Mathematics Pedagogical Goals (MPG) – can be emphasized in different ways for different ends. As per Gutstein (2006), SJPG include: developing positive social and cultural identities, reading the world with mathematics, and writing the world with mathematics (e.g. using mathematics to change the world). MPG include: reading the mathematical world (e.g. developing mathematical powers), succeeding academically in the traditional sense, and changing one's orientation towards mathematics. The last of these goals aims to help transition students from "seeing it [mathematics] as a series of disconnected, rote rules to be memorized and regurgitated, to a powerful and relevant tool for understanding complicated, real-world phenomena" (Gutstein, 2006, p.30).

While deeply interconnected, these goals can be difficult to negotiate and there is a tendency for teachers to prioritize one set over the other (Bartell, 2011/2013). Bartell observed that balanced learning goals of improving student understanding of world issues and of mathematics was not obvious to teachers, who were more likely to use social justice context problems to emphasize awareness of the social issue rather than develop mathematical knowledge. Indeed these teachers were observed to sacrifice mathematical content or develop lessons around previously acquired mathematical knowledge. In order to facilitate negotiation and balance of Mathematics and Social Justice Pedagogical Goals, more research is needed to identify and analyse teachers' understanding of social justice context problems, the related mathematical content, and their orientation toward that content.

Methodology

In this study, the TMSJ framework was used to develop and analyse social justice context problems related to the themes itemized at the beginning of this proposal. The specific tasks and data collection instruments will be provided in session. Briefly, participants were presented with a situation and asked to engage with some context problems which ranged from an explicit address of math concepts (e.g. represent data as graphs in different ways) to implicit suggestions (e.g. create the 'best' travel itinerary and justify your choices). Some further details are provided in the discussion section. Part of the broader research attends to issues in task design, but these are outside the scope of this proposal.

Participants for this study included 25 pre-service secondary school teachers enrolled in a methods course. Data collection occurred in multiple stages and used a mix of written response and one-on-one follow up interviews. Briefly, participants were asked to select from a variety of tasks that situated secondary school mathematics within issues of social justice. They had two months to engage with the task, which included in-class time for question-posing and collaboration. Following that, a one-on-one interview which invited them to reflect explicitly on the embedded mathematics, and their experiences engaging with the context problem, was administered. Data analysis is on-going, and preliminary results and trends are highlighted below.

Discussion

While there is not space to fully discuss the results, some highlights are presented. Analysis will focus on the MPG component to Gutstein's (2006) TMSJ framework. Preliminary results suggest that while participants saw a connection between the social justice issues and mathematics, and expressed a need to use mathematics to respond to the task, they had difficulty identifying what specific mathematical concepts were relevant.

For instance, in a question related to international fair trade practices, participants were asked to create a travel itinerary for their funding agency and to use mathematics and/or statistics to justify their choices. For the task designers the idea of optimization (of time, distances, money, etc.) was at the heart of this problem. However this was not clear for participants, even when (heavily) prompted. For example, Debbie noted that data management was involved in this problem, but she did not see a connection to optimization: "as far as optimization… I don't see how that would fit in… because when I think of optimization, I think of, I don't know, parabolas and things like that, and finding the optimum value." Debbie went on to identify issues connected with optimization – "dealing with money, dealing with schedules, dealing with time" – though she still did not see the connection herself. Interestingly, trends in the responses to this context problem included language which referred to optimization (e.g. choosing locations and means of transportation to minimize travel times and fees), but which was not perceived as optimization by participants whose attention tended to remain on the social justice issues related to fair trade. Connecting this observation to the MPG listed above, there appears to be a relationship amongst "reading the mathematical world" (Gutstein, 2006, p.24), changing one's orientation towards the subject (p.26), and *changing one's orientation toward the specific subject matter*. In particular for this example, participants needed to think about optimization in a broader context than finding vertex points of parabolas in order to notice it in a context about travel plans.

Participants had different difficulties with context problems that more explicitly identified the mathematics required to address the social justice issue. For example a context problem which asked participants to compare specific data related to cost of living and food subsidies in a major metropolis versus a remote location in their home country was seen as challenging because it contrasted with prior school experiences. Rudy explained that "school has created a kind of stereotype" but for this problem "there's no neat and final solution" because one cannot separate the data from its social context ("you have to see it all together," and "you need to learn about the social issues… otherwise you get confused with the data"). For Rudy, the task "reinforced how powerful mathematics is" and opened for him "a new window… to make teaching relevant, meaningful, helpful…" These comments exemplify the general reaction to this context problem. For the most part, participants described the power of using such problems for reinforcing math, but not to learn it. As Frank put it "though I do believe that it is important to help students see how math may connect to social issues in society… I do feel that it is necessary to keep to the curriculum… I also feel that that to invest the amount of time that this [context problem] would require… I would feel like I'm doing my students an injustice by not spending more time on the mathematics." Thus while Rudy's comments suggest a shift in orientation towards math teaching, Frank highlights the tensions and challenges of such a shift. It is interesting that Frank would use the word "injustice" to describe teaching mathematics through a social justice context problem, and this suggests a possible refinement of the TMSJ model when applied to teacher education.

Bartell (along with Gutierrez, 2009) notes that "learning to teach mathematics for social justice is a complex, long-term process and adequate contextualization of social issues, for example, will not occur in the course of one professional development experience" (2013, p.160). This comment can also apply to learning math from social justice, as noticing relevant mathematical ideas which may be embedded in a social context is not obvious, even for individuals with substantial mathematics backgrounds. Further, for teachers to want to teach in such a way, a shift in orientation may be necessary. Extending on Gutstein's (2006) model, there is a need for teacher educators to incorporate pedagogical goals that foster shifts in orientation towards (i) mathematics teaching, and (ii) achieving success in school mathematics. What might such shifts look like? How can teacher educators facilitate "seeing" math in social issues and then mobilizing?

References

Bartell, T. (2011). Learning to teach mathematics for social justice: Negotiating social justice and mathematical goals. *Journal of Research in Mathematics Education, 41*, 5-35. (reprinted 2013, *JRME*, *44*, Special Equity Issue, p.129-163)

Bateiha, S. (2010). *Mathematics for critical numeracy: A case study of social justice mathematics course for preservice elementary teachers*. Unpublished doctoral dissertation, University of Oklahoma, USA.

Beswick, K. (2011). Putting context in context: An examination of the evidence for the benefits of 'contextualised' tasks. *International Journal of Science and Mathematics Education, 9,* 367 – 390.

Esmonde, I., Blair, K. P., Goldman, S., Martin, L., Jimenez, O., Pea, R. (2013). Math I am: What we learn from stories that people tell about math in their lives. In B. Bevan, P. Bell, R. Stevens, & Razfar, A. (Eds.), *LOST Learning: Learning about Out-of-School-Time Learning Opportunities*. (pp. 7-28). New York: Springer.

Gutierrez, R. (2009). Enabling the practice of mathematics teachers in context: Toward a new equity research agenda. *Mathematical Thinking and Learning, 4*(2&3), 145-187.

Gutierrez, R. (2013). The socio-political turn in mathematics education. *Journal for Research in Mathematics, 44*, Special Equity Issue, 37 – 68.

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education, 34*(1), 37–73.

Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy of social justice.* New York, NY: Routledge.

Stinson, D.W. (2013). Negotiating the "white male math myth": African American male students and success in school mathematics. *Journal for Research in Mathematics, 44*, Special Equity Issue, 37 – 68.

93

An eye to the horizon: The case of Delia's Hexagon

Ami Mamolo
University of Ontario Institute of Technology

This paper explores pre-service secondary school mathematics teachers' preferences when advising a student on how to determine the area of an irregular hexagon. The research attends to participants' personal mathematical knowledge, as interpreted through the lens of Knowledge at the Mathematical Horizon. Philosophical notions of inner and outer horizons of conceptual objects are adapted to provide a refined analysis of participants' personal strategies and preferences as evoked by an unconventional problem. The interplay amongst participants' understanding of mathematical structure, their focus of attention when interpreting a problem, and the advice they offer to a student are of interest. Implications for teacher education and further avenues of research are suggested.

Key Words: Knowledge at the Mathematical Horizon; Pre-service Teacher Education; Geometry

Personal mathematical knowledge of teachers and its influence on lesson planning and implementation continue to draw attention from researchers at the undergraduate level (e.g. Watson and Chick, 2013). Facets of teachers' mathematical knowledge as directly related to the curriculum have been widely discussed, categorized, explored; and the research is well-known (e.g. Ball, Thames & Phelps, 2008). With the focus of these conversations in mind, eyes turn toward the horizon – namely, teachers' knowledge at the mathematical horizon (KMH) as introduced by Ball and Bass (2009) and extended by Zazkis and Mamolo (2011).

This research is part of a broader study which explores pre-service teachers' KMH as they respond to and prepare for teaching situations. This proposal explores the connection between participants' personal preferences for addressing a novel (for them) problem and their related expectations for student learning. Specifically: *What influences pre-service secondary mathematics teachers' preferences when considering recommendations for how to determine the area of an irregular hexagon, and what are the bases for these preferences?* The analysis intends to shed new light on how different facets of KMH may manifest in pre-service teachers' address of a teaching situation, and what consequences this might have for student learning.

Background

Mathematical knowledge for teaching, in teaching, of teachers, has been widely discussed from a variety of perspectives (e.g. Ball et al., 2008; Davis & Simmt, 2006). Much of the current discussion draws on Shulman's (1986) landmark distinction between Subject Matter Knowledge (SMK) and Pedagogical Content Knowledge (PCK), refining and extending understanding of what knowledge is required for the tasks of teaching. While the debate continues, there seems to be consensus that teachers "need to know more advanced mathematics than the mathematics they are teaching" (Mason & Davis, 2013, p.194). The nature of such knowledge also attracts interest, and as Rowland and Zazkis (2013) suggest, "one's stance on the mathematical knowledge needed (or essential) for teaching depends on one's perception of teaching itself" (p.138). In their perspective, teaching involves dealing with the unanticipated – e.g. taking advantage of unexpected opportunities for making connections or extending student thinking. They therefore state that "mathematical knowledge beyond the immediate curricular prescription is beneficial and demonstrably essential" (*ibid.*).

In Ball et al.'s (2008) well-known refinement of SMK and PCK, they note that the SMK required for teaching are "knowledge and skills not typically taught to teachers in the course of their formal mathematics preparations" (p.402). They further state that "Teachers who do not themselves know a subject well are not likely to have the [pedagogical content] knowledge they need to help students learn this content" (p.404). Similarly, Potari et al. (2007) observed that robust subject matter knowledge allows teachers to interpret and develop student ideas with greater ease and effectiveness. They also suggest that teachers' ability to connect different mathematical areas and their awareness of the relevance of these connections were part and parcel to their ability to effectively create a rich mathematical learning environment. Ball, Lubienski, and Mewborn (2001) highlighted that: "It is not only what mathematics teachers know but also how they know it, and what they are able to mobilize" (p.451).

Extending on these studies, this research considers pre-service teachers' address of a student's attempt to solve a non-routine problem, which can be addressed in a variety of ways with connections to core concepts in secondary school curricula. Participants' SMK is analysed via the sub-category Knowledge at the Mathematical Horizon, which is discussed in the following section. The intent is to offer a refined look at how this specific aspect of individuals' mathematical knowledge can manifest in, and influence, teaching situations.

Theoretical Framework

Knowledge at the Mathematical Horizon (KMH) is described as a structural, connected, and robust understanding of mathematics that goes beyond what is taught in school curricula. Ball and Bass (2009; also Jakobsen, Thames, & Ribeiro, 2013) present KMH as a teacher's knowledge of students' horizon; it includes teachers' knowledge of major disciplinary ideas and structures, key mathematical practices, and core mathematical values as they relate to students' past and future learning (p.5). Other perspectives identify similar features, but focus on the *teachers'* horizon (e.g. Zazkis & Mamolo, 2011). This study also attends to teachers' horizon and what mathematics lies "in and out of focus" as they consider a hypothetical teaching situation.

In line with the description of horizon knowledge as connected, robust, and beyond school curricula, Zazkis and Mamolo (2011) extend the construct of KMH to focus on teachers' horizon by connecting it to Husserl's philosophical notion of a (conceptual) object's horizon. Husserl's description relates to an individual's focus of attention – in particular, when an individual attends to an object, the focus of attention centers on the object itself, while the 'rest of the world' lies in the periphery (Follesdal, 2003). With this perspective, an individual's KMH is contingent on the specific mathematical object under consideration – how it is understood, what aspects lie in focus or in the periphery, and what connections the individual is able to make between the in-focus and peripheral facets. What lies in the periphery is understood as the object's horizon, and according to Husserl it can be described by an *inner* and an *outer* horizon.

Inner horizon refers to specific attributes of an object which are not (at that moment in time) in focus for the individual. For example, if one were to imagine a hexagon, there are several attributes which might appear in focus – the number of edges and vertices, the lengths of the sides, etc. Other specific attributes of that hexagon – its lines of symmetry, area, etc. – lie in the periphery, and as such are elements of the inner horizon. There is a reflexive relationship between what lies in focus and in the inner horizon, and it depends on what catches (and keeps) the attention of the individual. An object's outer horizon refers to the "broader world" in which the object exists, and thus does not depend on the individual's focus of attention. The outer horizon includes features which embed the object in a greater structure, and consists of generalities exemplified by the specific object. For example, the fact that one may express measurements of the hexagon via algebraic equations (such as for perimeter or size of interior angles) would lie in the outer horizon, exemplifying structural connections between strands (e.g. algebra and geometry) and between concepts (e.g. ratios of lengths and angles).

With this in mind, a teacher's KMH can be interpreted as knowledge of mathematical objects' inner and outer horizons. Zazkis and Mamolo (2011) use this construct to explore examples of KMH as it influenced teachers' decisions in classroom situations. Extending on this work, this research explores what mathematical knowledge is "in focus" as pre-service mathematics teachers addressed a non-routine problem about the area of an irregular hexagon. The analysis attends to specific instances of inner and outer horizon knowledge and how this knowledge influenced participants' preferred recommendations for a hypothetical student.

Methodology

Data collection occurred in two stages via written questionnaire with 20 pre-service secondary mathematics teachers enrolled in a teacher education program. The questionnaires were administered one week apart and took approximately 30 minutes to complete. Participants were informed of the scope of the questionnaires, which sought to explore their mathematical and pedagogical knowledge given a hypothetical situation. They were not told of the specific content in advance, aside from the fact that the second questionnaire would follow-up on ideas raised in the first session. Participants were told to answer honestly and reflectively, and that it was okay to say "I don't know." It was also emphasized that there was no "right answer."

The first questionnaire, depicted in Figure 1 below, was designed to uncover participants' strategies and approaches when advising a student. "Delia's hexagon" was chosen for this questionnaire because of the many applicable concepts and strategies which would (i) allow participants to solve by various means, and (ii) provide information on what relevant (and irrelevant) concepts were evoked and remained "in focus." During this stage of the research, a diagram of Delia's hexagon was deliberately omitted since how (and whether) participants constructed their own diagrams would provide further insight into their thinking.

> Imagine you are a teacher in the following situation: Delia, a high school student with good grades, is working on an extra-curricular math problem and approaches you for help. Here is the problem:
>
> You are given a hexagon ABCDEF, where the lengths of the sides are equal to AB = CD = EF = 1 and BC = DE = FA = $\sqrt{3}$, and AB is parallel to DE, BC parallel to EF, and CD parallel to FA.
>
> 1. What is the measure of each interior angle?
> 2. What is the area of the hexagon?
>
> Delia has found that all of the interior angles are of equal measure, but is unsure how to find the area. *How do you recommend Delia go about finding the area?*

Figure 1: First Questionnaire: Introducing Delia's Hexagon

Since the second questionnaire was designed as a follow-up, it is appropriate to quickly summarize the trends and initial analysis of the first questionnaire before presenting the task. Briefly, of the 20 participants 18 drew diagrams and all 20 described "deconstructing" the hexagon into smaller "easier" shapes. Fifteen participants drew regular hexagons, with the most common diagrams depicted in Figure 2 below. The two most prominent trends in participants' recommendations for Delia were: (i) based on broad ideas, such as "put in lines to break up the hexagon into shapes which we have established rules and laws to work with" (Sophia); and (ii) giving step-by-step procedures of how to solve, such as “she can solve the area of the two triangles and the rectangle in the middle using formulas for the areas. Once this is calculated, she can just add the area of the rectangle and the two triangles” (Abigail).

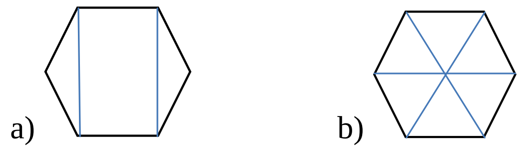


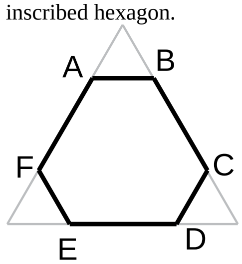
Figure 2: Common “dissecting strategies”

Initial analyses suggested participants over-relied on the regularity of their depicted hexagons, using strategies that were either inappropriate (e.g. Fig.2a) or incomplete (e.g. Fig.2b) to generalize to the irregular case. Thus, the second questionnaire (see Fig.3 below) included recommendations that (i) had diagrams of a regular and irregular hexagon, (ii) could apply in general without introducing any additional mathematics, and (iii) reflected and contrasted participants’ inclination to dissect the hexagon. In the following section, participants’ responses to the second questionnaire are analysed in depth. Due to space limitations, the focus is on identifying specific instances of participants’ KMH via the refined lens of inner and outer horizons, which exemplified the trends and themes observed more generally in the data.

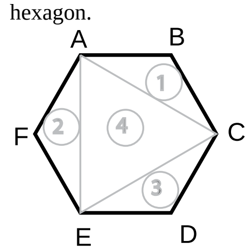
Recall Delia’s hexagon ABCDEF, with sides lengths AB = CD = EF = 1 and BC = DE = FA = $\sqrt{3}$.

To determine the area, Delia was given a variety of different recommendations. Here are two of them:

Recommendation A:
Extend the hexagon into an equilateral triangle as in the figure below. Then use the areas of the large triangle, and small outer triangles, to determine the area of the inscribed hexagon.



Recommendation B:
Decompose the hexagon into three triangles (1, 2, 3, which are all equal), and an equilateral triangle 4, as in the figure below. Then sum the areas of the inscribed triangles to determine the area of the hexagon.



Which approach do you prefer, and why?

Figure 3: Second Questionnaire: Participant Preferences

Results and Analysis

Trends observed in response to the second questionnaire:

1. Participants who preferred rec. A (11 out of 20) attended to structural features and consequences of the provided diagram;
2. Participants who preferred rec. B (9 out of 20) attended to surface features of the solving process and their prior personal experiences;

With respect to the first trend, attention to structural features and consequences of the diagrams was noted both in participants’ acceptance and critique of the two recommendations, respectively. For example, in Sarah’s response she identified the area of the hexagon as the difference between the areas of the large triangle and three smaller ones in rec. A, which can be interpreted as an instance of KMH at the outer horizon – that a shape's measurements may be determined indirectly through knowledge of other shapes can be considered as knowledge of a broader mathematical structure. Sarah explained rec. A was "easier" because "you're only using equilateral triangles," a fact which she explained verbally as well as diagrammatically (see Fig.4). In her reasoning she made use of structural properties of equilateral triangles, as depicted in Fig.4, and used these to deduce information about Delia's hexagon (outer horizon knowledge). She also contrasted this approach with rec. B, which she wrote "is unclear and... seems as though there is a lot more work to finding the areas of the triangles." This may be seen as an awareness that some of the important structural features of rec. A were not present in B.

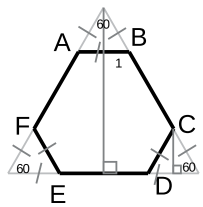


Figure 4: Sarah's diagrammatic reasoning

Sarah's focus on the triangles further suggests a relationship between an individual's inner and outer horizon knowledge. Her outer horizon knowledge of inferring from shapes allowed her to shift Delia's hexagon to the periphery of her attention, and as such it became part of the triangle's inner horizon (since it is specific to this triangle). This shift permitted Sarah to reason with what she believed was "easier," and it influenced how she would respond to Delia.

Similarly, Miles also attended to the specific structure of the provided diagrams, showing evidence of his inner horizon knowledge. He wrote: "It must also be noted that both figures shown are only one possible configuration. In fact, figure B is further from an accurate scale representation than figure A." Miles was one of four participants who referred to the inaccuracy of rec. B's diagram, and it was clear that an awareness of different hexagonal configurations remained in mind as he assessed the general applicability of the two recommendations. He noted that even "if the internal angles aren't equal, figure A's approach can still be used. The triangle form though, may not be equilateral, but it will be isosceles." In contrast to prior research which observed that teachers' images of hexagons tended to be restricted to regular prototypes (e.g. Ward, 2004), Miles was considering hexagons more broadly. His response suggests that he was able to reason with these shapes without having them directly in view. His consideration of how the encompassing triangle would differ depending on the specific hexagon and of how "A's approach can still be used" more generally, instantiate both inner and outer horizons, respectively. Both Miles and Sarah also noted that both recommendations should be shown to Delia since it will "deepen [her] understanding of the concepts involved."

This latter reflection on the value of both approaches is worth noting as it contrasted with common responses that preferred rec. B as more familiar, comfortable, and more closely connected to strategies used in school. The dominating factors for these participants were familiarity and personal comfort levels, and these influenced both what was deemed appropriate for Delia and what seemed to be "allowed in view" for participants. For example, Abigail claimed that "rec. B is the approach I would take because of the way I learned geometry. The hexagon divided into triangles is the approach I learned in school." She went on to say that the "subtraction method [is] confusing to me, but adding small shapes to make a big shape is easy." Abigail's desire to stick to the approach she learned in school suggests a limited KMH, and she did not seem to find beneficial "mathematical knowledge beyond the immediate curricular prescription" (in contrast to Rowland and Zazkis, 2013). Further, her reluctance to consider an approach that lay outsider her prior experience also seemed to play a part in directing her attention toward superficial features of the recommendations – such as the relative sizes of the shapes in question, and the challenges of adding versus subtracting. These ideas were echoed by the majority of participants who preferred rec. B.

For instance, Victor replied: "I find it easier to conceptualize adding small portions to get the new total portion instead of calculating a larger portion and subtracting from it… only one function is required (addition) vs. [rec. A] which requires two functions (addition and subtraction)." In all cases, these operations were not considered in light of the details of their implementation – that is, no one attended to what was going to be added or subtracted, or to how difficult it would be to determine these values (or how many "functions it would require"). Focusing on such surface features can be interpreted as limited horizon knowledge, as well as a resistance to consider deeply an approach that was outside of familiar repertoire. The two are connected, and in turn connect to how teachers may view and guide their students' learning. In contrast, with a robust KMH, and a willingness to apply it, individuals could have analysed the level of difficulty of the arithmetic with respect to the specific features of Delia's hexagon (inner horizon), and provided justification for their preferences that spoke to the general applicability of the proposed recommendation (outer horizon).

Concluding Remarks

What influences pre-service secondary mathematics teachers' preferences when considering recommendations for how to determine the area of an irregular hexagon, and what are the bases for these preferences? Returning to this research question, several factors were found. In resonance with research done with children (e.g. Clements & Battista, 1992; Walcott et al., 2009), participants relied heavily on prototypes of regular hexagons, which was surprising given their strong mathematics backgrounds. Many participants cited surface features as the bases for their preferences, attending to relative sizes of shapes as well as the number of operations needed as indicators of the levels of difficulty of solving strategies. In these cases, specific features of the recommendations were ignored in favour of general observations and personal preferences. These considerations resulted in responses that were either inappropriate (such as in Fig.2) or misleading (e.g. Abigail and Victor), and were interpreted as illustrating limited KMH.

Although Delia's hexagon was new for all participants, some were more inclined to make connections beyond their prior school experiences and to consider the problem on a deeper level. These participants (e.g. Sarah and Miles) attended to structural features of the mathematics, were more flexible with what was in- and out- of focus, and were interpreted as demonstrating more robust KMH, as it related to both inner and outer horizons of the mathematical entities in question. The analysis suggests that comfort and flexibility in what horizon knowledge is accessed and mobilized are important features in assessing and recommending appropriate solving strategies for learners, and it makes a case for developing such flexibility in teacher education programs. Further, participants who demonstrated this flexibility also seemed more willing to set aside their personal and initial preferences for advising Delia, and this is particularly significant when considering the potential impact of a teacher's response to a student's unconventional or unexpected approaches. These results suggest a need for more research into how limited or robust Knowledge at the Mathematical Horizon, with particular reference to their personal mathematical knowledge (e.g. inner and outer horizons), can influence how teachers' interpret, predict, and respond to student thinking.

References

Ball, D.L. & Bass, H. (2009). With an eye on the mathematical horizon: knowing mathematics for teaching to learners' mathematical futures. Paper presented at the 43rd Jahrestagung der Gelleschaft fur Didaktic der Mathematik, Oldenburg, Germany.

Ball, D.L., Lubienski, S.T., & Mewborn, D.S. (2001). Research on teaching mathematics: The unresolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of Research on Teaching* (pp. 433–456). Washington: American Educational Research Association.

Ball, D., Thames, H. M. & Phelps, G. (2008). Content Knowledge for Teaching. *Journal of Teacher Education, 59*(5), 389-407.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*, (pp. 420-464). New York, NY, England: Macmillan Publishing Co, Inc.

Davis, B. & Simmt, E. (2006). Mathematics-for-teaching: an ongoing investigation of the mathematics that teachers (need to) know. *Educational Studies in Mathematics, 61*(3), 293-319.

Follesdal, D. (2003). Husserl, Edmund (1859-1938). In Craig, E. (Ed*.) Routledge Encyclopaedia of Philosophy* (2nd edition, online). London, UK: Routledge. Retrieved July 19, 2010, from www.rep.routledge.com/article/DD029

Jakobsen, A., Thames, M.H., & Ribeiro, C.M. (2013). Delineating issues related to horizon content knowledge for mathematics teaching. Paper presented at the *Eighth Congress of European Research in Mathematics Education* (CERME-8). Retrieved from: http://cerme8.metu.edu.tr/wgpapers/WG17/WG17_Jakobsen_Thames_Ribeiro.pdf

Mason, J. & Davis, B. (2013). The importance of teachers' mathematical awareness for in-the-moment pedagogy. *Canadian Journal of Science, Mathematics and Technology Education, 13*(2), 182-197.

Potari, D., Zachariades, T., Christou, C., Kyriazis, G. & Pitta-Pantazi, D. (2007). Teachers' mathematical knowledge and pedagogical practices in the teaching of derivative. In Pitta–Pantazi, D. & Philippou, G. (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 1955 - 1964).

Rowland, T., & Zazkis, R. (2013). Contingency in the mathematics classroom: Opportunities taken and opportunities missed. *Canadian Journal of Science, Mathematics and Technology Education, 13*(2), 137-153.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Walcott, C., Mohr, D., Kastberg, E. S. (2009). Making sense of shape: An analysis of children's written responses. *The Journal of Mathematical Behavior, 28*, 30-40.

Ward, R. (2004). An Investigation of K-8 Preservice Teachers' Concept Images and Mathematical Definitions of the polygons. *Issues in Teacher Education, 13,* 39- 56.

Watson, A. & Chick, H. (Eds.). (2013). Personal Mathematical Knowledge in the Work of Teaching [Special Issue]. *Canadian Journal of Science, Mathematics and Technology Education, 13*(2).

Zazkis, R. & Mamolo, A. (2011). Reconceptualizing knowledge at the mathematical horizon. *For the learning of Mathematics, 31*(2), 8-13.

A FRAMEWORK AND A STUDY TO CHARACTERIZE A TEACHER'S GOALS FOR STUDENT LEARNING

Frank S. Marfai
Arizona State University

In this study, a secondary school teacher's goals for student learning were characterized using a framework that emerged from prior work. Observed lessons spanning the use of both conceptually rich and conceptually poor curricula were analyzed and lead to unexpected findings, suggesting that both challenges and opportunities for professional development endeavors exist that center around perturbing a teacher's goals.

Key words: Teacher Goals, Mathematical Knowledge for Teaching, Teacher Knowledge, Teacher Beliefs, Professional Development

It is widely known that mathematics teaching in the United States has been characterized as procedural and disconnected (Ma, 1999; Stigler & Hiebert, 1999), with little focus on understanding how mathematical concepts develop and how they are connected. In recent work it has also been documented that it is common for teachers to teach in a manner in which they were instructed as students, and that making the transition to value conceptual learning and teaching is a difficult transition for teachers to make (Sowder, 2007).

Theoretical Framework

Mathematical knowledge for teaching (MKT) has been described as the domains of knowledge that include a teacher's subject matter knowledge and her pedagogical content knowledge (Ball, 1990; Hill, Ball, & Schilling, 2008). MKT has also been described as a teacher's key developmental understandings and how they influence a teacher's practice (Silverman & Thompson, 2008). It has been reported that many teachers do not possess key developmental understandings of central ideas of secondary mathematics, and that these understandings can only emerge from experiences that promote perturbations that result in self-reflection.

A teacher's mathematical teaching orientation influences her classroom practices (Thompson, Philipp, Thompson, & Boyd, 1994). A teacher with a calculational orientation has an image of mathematics as an application of rules and procedures for finding numerical answers to problems. A teacher having a conceptual orientation has an image of mathematics as a network of ideas and relationships among these ideas, and strives to support students in developing coherent meanings among these ideas. I conjecture that a teacher's mathematical orientation is influenced by her MKT and that this knowledge may also impact the goals a teacher has for her students' learning and her teaching.

I will define a teacher's goal as a mental representation of what a teacher is trying to accomplish. This is similar to how other researchers (Locke & Latham, 2002) have categorized goals, although this perspective does not explain possible purposes or reasons why a teacher may pursue a goal (Pintrich, 2000). Research has shown that a teacher's goals for student learning do influence her development of powerful pedagogical content knowledge (Webb, 2011).

In the both studies on which this preliminary research report is based, I used Silverman and Thompson's (2008) construct of MKT as a lens for examining how a teacher understands ideas and connections among ideas, and how this influences her pedagogical decisions and actions. I hypothesize that the transformation of a teacher's key developmental understandings (Simon, 2006) into MKT is developmental as a teacher's orientation shifts from calculational to conceptual, and examining teachers' pedagogical goals for a lesson can lead to insights underlying this process of growth.

Research Question

My research question is as follows: How might a teacher's pedagogical goals for student learning be characterized in the context of using a curriculum promoting a conceptual orientation of mathematics, and how are they similar or different than when using a curriculum that promotes a calculational orientation?

Methods

To characterize a teacher's goals for student learning, two studies were conducted. The first study was a pilot project in which an initial goal framework emerged from two teachers' goals for student learning using grounded theory (Strauss & Corbin, 1990) keeping in mind Silverman and Thompson's characterization of MKT. In the second study (the follow-up to the pilot study), Robert (pseudonym) from Salt Valley High School (pseudonym) in a Southwestern state was selected for observation during two chapters in which he taught Trigonometry during the Spring 2013 semester. Robert was teaching Precalculus for the third time using the same conceptually rich curriculum as the teachers in the first study, although he had supplemented the course with materials from a traditional textbook. Robert has been teaching for 13 years total at the same high school. Robert was identified as a teacher whose key developmental understandings of the Precalculus curriculum were well connected and whose pedagogical actions indicated an inclination to act on student thinking.

Twenty-nine classroom observations were videotaped that primarily covered two chapters from different texts focusing on trigonometry, in particular angle measure, trigonometric functions, identities, and applications using trigonometric functions. At the end of class, the researcher gave a short questionnaire that included asking about Robert's instructional goals, and his goals for student learning that day; he answered the questionnaire on the same day via email. Robert's goals for student learning were then coded using the framework developed from the pilot study. The chapters under which the observations were performed had a conceptually rich chapter from reform oriented curricular materials, followed by a conceptually poor chapter and sections of a traditional textbook. In addition to observing how curricular context affected Robert's goals for student learning, the researcher also tested the stability of Robert's goals through follow-up questions designed to perturb his goals to higher levels in the framework.

Results

Robert's goals were coded using the goals framework developed from the pilot study; it consists of seven levels (rated from 0 to 6 – representing a spectrum of product focused actions to do at lower levels in the framework, towards goals focused on student thinking and ways to support such thinking at higher levels in the framework) of teacher's goals for student learning (TGSL) at increasing degrees of sophistication. Table 1 contains the framework used in this study.

Table 1. Levels in a Teacher's Goals for Student Learning

Goal Coding	Description
TGSL0	Goals are not stated, or the teacher states that the goals of the lesson are unknown.
TGSL1	Goals are a list of topics that a teacher wants her students to learn in the lesson, each associated with an overarching action.
TGSL2	Goals are a list of topics that a teacher wants her students to learn in the

	lesson, each associated with a specific action.
TGSL3	Goals are doing methods of mathematics that a teacher wants her students to learn in the lesson.
TGSL4	Goals are getting students to think about the mathematics in the lesson, without the ways of thinking articulated.
TGSL5	Goals are getting students to think about the mathematics in certain ways during the lesson.
TGSL6	Goals are about developing ways of thinking about the mathematics in the lesson, with attention to how that thinking may develop.

After coding Robert's goals, statements of his goals for student learning ranged from levels 1 to 5 (see Table 2, below), even though based on classroom observations, in a majority of class sessions Robert made pedagogical moves to model student thinking and he made decisions to act on his model of student thinking either at the group level or in a whole class discussion, with varying levels of success. Robert's pedagogical moves in more successful interactions initially suggested that he was mindful of student thinking in the planning process, and thus the reason why TGSL6 exists in the framework; however they appeared to be unstated goals. Goals rated at TGSL6 were accessible to Robert, however were only stated explicitly in limited contexts after moves were made by the researcher to perturb his goals.

Table 2. Teacher's Goals for Student Learning - Count (Percentage)

Goal Level	Conceptually Rich Curriculum	Conceptually Poor Curriculum
TGSL0	0 (0.0%)	0 (0.0%)
TGSL1	7 (17.1%)	2 (9.5%)
TGSL2	10 (24.4%)	1 (4.8%)
TGSL3	3 (7.3%)	8 (38.1%)
TGSL4	17 (41.5%)	7 (33.3%)
TGSL5	4 (9.8%)	3 (14.3%)
TGSL6	0 (0.0%)	0 (0.0%)
All Stated Goals	41 (100.0%)	21 (100.0%)

Another interesting finding is when looking at Robert's goals for student learning alone, there was no way to distinguish between the modalities of classes that consisted mostly of group work, meaning making, and class discussion versus those lessons in which direct instruction with limited opportunities for group work were observed; Robert's goals for student learning were indistinguishable when using reform-oriented curricular materials were used, versus when a traditional textbook was used.

Questions

The two primary questions I have for audience members in which I would value the feedback and comments from the research community are as follows.

In early analysis, I have found that Robert's familiarity with the curriculum resulted in superficial attention to planning. Therefore I hypothesize that explicit attention to goals on ways students' thinking could be developed, promoted, or supported (TGSL6) are not part of Robert's natural inclinations. The goals framework has been used as a tool to characterize a teacher's goals for student learning, and I have used it as a tool for professional development (to promote self-reflection through follow-up question to move a teacher's goals towards higher levels in the framework). Moves by the researcher to cause perturbations were met with limited success occurring only in certain contexts. My first question is how might a researcher help shift a teacher's natural disposition toward goals for student learning that would be rated as TGSL6, and what role may a teacher's beliefs and his or her mathematical knowledge for teaching contribute to resistance, or amenability, to such shifts?

I conjecture that a teacher's planning process and mathematical knowledge for teaching contribute to her or his goals supportive of student autonomy and reflection in learning of mathematical ideas, through goals that promote students making mathematically relevant conjectures and reflections. From the analysis of the data, I did not see explicit evidence of such goals for student learning for Robert, and his classroom practice did not reflect consistent attention to these goals either, regardless of the type of curriculum used. My second question for the research community is how might teacher's goals for student learning be perturbed to include opportunities for student conjecture and reflection in order to promote student autonomy?

References

Ball, D. L. (1990). The Mathematical Understandings That Prospective Teachers Bring to Teacher Education. *The Elementary School Journal, 90*(4), 449-466.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education, 39*(4), 372-400.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey. *American Psychologist, 57*(9), 705–717.

Ma, L. (1999). *Knowing and Teaching Elementary Mathematics*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Pintrich, P. R. (2000). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology, 25*(1), 92-104.

Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education, 11*, 499-511.

Simon, M. A. (2006). Key Developmental Understandings in Mathematics: A Direction for Investigating and Establishing Learning Goals. *Mathematical Thinking and Learning, 8*(4), 359-371.

Sowder, J. T. (2007). The Mathematical Education and Development of Teachers. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 157-224). Charlote, NC: Information Age Publishing.

Stigler, J. W., & Hiebert, J. (1999). *The Teaching Gap*. New York: Free Press.

Strauss, A. C., & Corbin, J. M. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*: Sage Publications.

Thompson, A. G., Philipp, R. A., Thompson, P. W., & Boyd, B. A. (1994). Calculational and Conceptual Orientations in Teaching Mathematics. In A. F. Coxford (Ed.), *1994 Yearbook of the NCTM* (pp. 79-92). Reston, VA: NCTM.

Webb, M. (2011). *Teachers' Knowledge Is Built Around Their Goals.* In T. d. S. Lamberg & L. R. Wiest (Eds.), Proceedings of the Thirty-third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Nevada: University of Nevada, Reno.

AN INVESTIGATION INTO STUDENTS' USE OF GIVEN HYPOTHESES WHEN PROVING

Kathleen Melhuish
Portland State University

The mathematical practice of strengthening or weakening a theorem requires careful attention to hypothesis and conclusion. Selden and Selden (1987) reported that students often unintentionally weaken theorems raising concerns of undergraduates' attention to hypothesis. In this paper, I consider both the prevalence of this error and what the practice of strengthening/weakening a theorem may look like. A survey of prove/disprove prompts was piloted with five graduate students. A subset of these prompts was then given to undergraduates in an introductory group theory course. Preliminary results indicate that the error of weakening the theorem is prevalent amongst both populations. The graduate students participated in follow-up interviews where they were prompted to strengthen/weaken conjectures to further examine their attention to the hypotheses. In this preliminary report, I will present the survey results and one graduate case to illustrate what the practice of strengthening/weakening a theorem may look like.

Key words: proofs, group theory, hypotheses

Selden and Selden (1987) have previously presented a large taxonomy of proof errors, many of which have not been researched further. Of these error-types, several involved students making unfounded assumptions such as: real numbers laws are universal and weakening the theorem. In order to explore this phenomenon further, a survey was created using prompts were false assumptions could easily lead to an invalid proof.

In mathematics, the practice of strengthening or weakening conjectures is standard (Pólya, 1990). Yet, in traditional undergraduate classrooms, students are rarely asked to engage in this practice. Rather, students often start with a statement known to be true and begin their exercise of proving. When students unintentionally make a hypothesis weaker (by assuming properties not given) or fail to use all pieces of a hypothesis (essentially proving a stronger statement), they show an unawareness of what system they are in fact working in and likely lack the tools to engage in the practice of strengthening or weakening conjectures appropriately.

In this report, I will be considering the following research questions: 1. Is the error of weakening the theorem a frequent issue in introductory group theory courses? 2. Are students aware of what parts of a hypothesis are being used?

Background

A large body of knowledge has been developed in the past years concerning students struggles to prove statements, analyze proofs or evaluate conjectures (Dreyfus, 1999; Harel & Sowder, 1998; Hart, 1994; Selden & Selden, 1987, 2008; Weber, 2001). Selden and Selden (1987) note students often will weaken the theorem during the proving process. They define weakening the theorem as, "when what is used [to prove] is stronger than the hypothesis or when what is proved is weaker than the conclusion" (p. 10). This occurs when a student assumes a mapping is bijective or that a that a semigroup is actually a group.

When a mathematician is aware of this weakening or strengthening of a conjecture, they are engaged in the process of generalizing or specializing (Pólya, 1990). Pólya defines *generalization* as "passing from the consideration of a given set of objects to that of a larger set, containing the given one" (p. 12) whereas *specialization* is "passing from the consideration of a given set of objects to that of a smaller set, contained in the given one" (p.

12). Consider the conjecture: all 1-1 group homomorphisms preserve the abelian property. The statement, all group isomorphisms preserve the abelian property would be a specialization whereas all group homomorphisms preserve the abelian property would be a generalization.

Specialization and generalization require an awareness of the hypothesis and what properties apply to the set of objects being considered. In this report, I will present data showing students often create inappropriate proofs indicating a failure to pay attention to these attributes. I will also consider how graduate students deal with hypotheses.

Methods

The results reported here come from two early phases of this study. A series of questions were piloted with five graduate students (two current PhD students in Mathematics Education, two planning to continue into the PhD program and one current Masters of Science in Teaching Mathematics student). Initially, the five students took home a survey containing six prove/disprove prompts. After the initial exploration, each student participated in a semi-structured follow-up interview.

For the purpose of this proposal, I will concentrate on the two prompts found in Table 1. The two prompts are altered slightly for the two survey versions. The intent was for each survey to contain a statement that was too weak, and a statement that was too strong. During the follow-up interview, each of the students was asked to consider if a statement could be weakened (if false to make true) or if a statement could be strengthened (if true.) For this preliminary report, one of these cases will be presented in detail.

The second phase of the study consisted of surveying an introductory group theory class. Version A and B of the survey were distributed as an extra credit homework assignment. 19 undergraduate students responded (9 with version A and 10 with version B). The results of the survey will be broken down into categories based on their correctness, and usage of hypotheses.

Table 1
Survey Prompts

1. Let ϕ be a 1-1 homomorphism from $(G,\circ)$ to $(H,*)$. If G is an abelian group, then H is an abelian group. *Version A*: Prompt 1
1. Let ϕ be an isomorphism from $(G,\circ)$ to $(H,*)$. If G is an abelian group, then H is an abelian group. *Version B*: Prompt 1
2. Let f and g be group homomorphisms from abelian group $(G,\circ)$ to abelian group $(H,*)$. Define $h: G \to H$ by $h(x) = f(x) * g(x)$. Then h is also a homomorphism. *Version A*: Prompt 2
2. Let f and g be group homomorphisms from $(G,\circ)$ to $(H,*)$. Define $h: G \to H$ by $h(x) = f(x) * g(x)$. Then h is also a homomorphism. *Version B:* Prompt 2

Results

The five graduate students surveyed provided interesting cases where each student engaged with prompts in very different matter. During this initial investigation hypothesized errors came to fruition including one student utilizing the commutative property when a given group was not assured to be abelian and two students missing the necessity of a homomorphism being surjective to guarantee commutativity is preserved. The following case presents one way this generalization/specialization process could unfold.

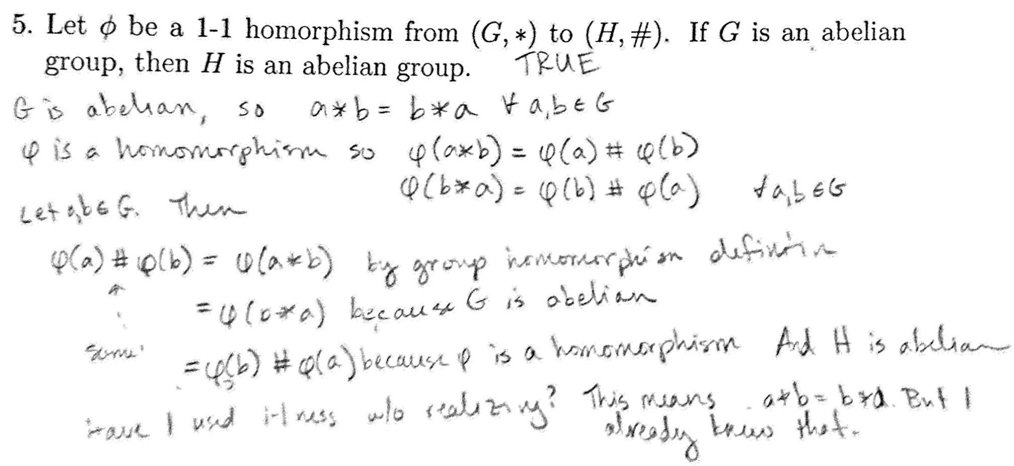


Figure 1. Carissa's Survey Response.

Case: Carissa. Carissa's survey response can be seen in Figure 1. She began her proof by writing the definition of abelian and homomorphism. Carissa then chose elements from G and argued about the commutativity of their image. (This would be a valid argument if the mapping was surjective and all elements in H were images of elements in G.) Carissa noted that she did not use 1-1.

During the follow-up interview, I prompted Carissa to expand on her 1-1 statement. Carissa explicitly stated concern: "Just in general terms, you should use all of your given information when you are proving something." This attention to the hypothesis was unique amongst her peers and also represented the view that the strongest version of a theorem is the standard.

Carissa then walked me through her argument confirming her false assumption. "I started with elements in H. Elements in H are phi of elements in G and I want to show that ends up being the same as pound sign the two elements backwards." Carissa then momentarily considered if 1-1 is what allowed these elements to "exist." I prompted Carissa to share her definition of 1-1 at which point she started drawing the diagram seen in Figure 3. She used the nonexample $y=x^2$ to explain that when a mapping is 1-1. "I'm figuring out my definition based on what it doesn't do. 1-1 is when they don't go to the same thing. If I had this over here, I would know that there is one and only one that it went to." At this point, Carissa appears to be confounding 1-1 with bijective.

I then asked Carissa if she was making use of the 1-1. After looking at her diagram she responded, "It may be onto-ness not one to one ness. It should be more like: if I had an *a* and a *b* in H. So maybe this is not necessarily true because what I need is for there to be something over here." By diagraming and considering the definition of 1-1, Carissa realized her original mistake. When prompted to prove that it was not true, she sketched the diagram in Figure 2. Carissa drew one element in the domain and several elements in the range. She then made the argument specific saying, “Take a nice little identity group going to the symmetry group and just map the identity to the identity.” She pondered if this would be a homomorphism, argued quickly about identities operating on themselves, and declared “this is boringly working.”

Carissa’s survey result confirmed the expected error: treating the mapping as if surjective. Through exploration of the hypotheses, Carissa corrected her error realizing the homomorphism must be onto for the statement to hold. Carissa engaged in both specializing (surjective homorphisms) and generalizing (removing the one-to-one requirement.)

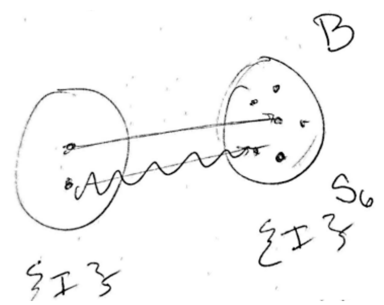


Figure 2. Carissa’s nonexample of a surjective homomorphism

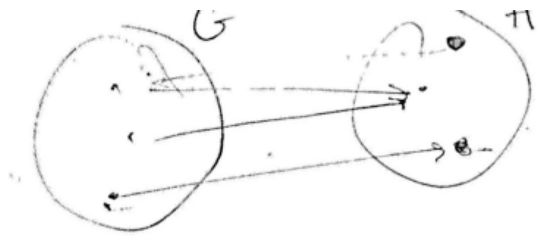


Figure 3. Carissa's nonexample of 1-1 mapping

Undergraduate survey results

Prompt 1. Version A. Of the 9 surveys, seven students incorrectly identified this statement as true. Each of the seven students picked elements in the image of mapping as opposed to beginning in H. (This would require the mapping to be surjective to be valid.) The two students who said the statement was false did not come up with a valid counterexample or note the need for onto.

Prompt 1. Version B. Of the 10 surveys, all students correctly identified this statement as true. Four students presented invalid or informal proofs noting that isomorphism preserves properties. Four students provided proofs beginning with image as done in version A while only two students currently utilized the onto nature of isomorphism.

Prompt 2. Version A. Of the 9 surveys, 7 students attempted this prompt. These seven students correctly identified the statement as true. In four of the proofs, the students correctly utilized H being abelian. The other three did not explicitly use H being abelian.

Prompt 2. Version B. Of the 10 surveys, 9 students attempted this prompt. Four of the nine marked the statement as true with three of those students assuming H was abelian. Of five false answers, two presented valid counterexamples and three noted that H needed to be abelian.

Sample student work can be seen in Figures 4 and 5 showing students unintentionally weakening the theorem. These samples representative of many of the students regardless of prompt version A or B.

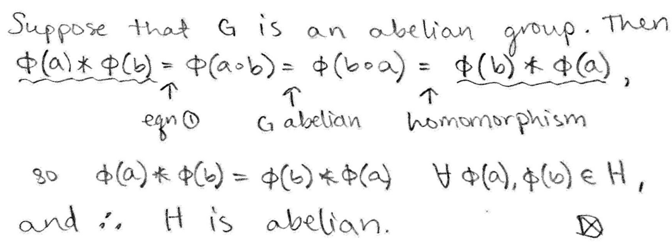


Figure 4. Student Response (Prompt 1, Version B: Homorphism)

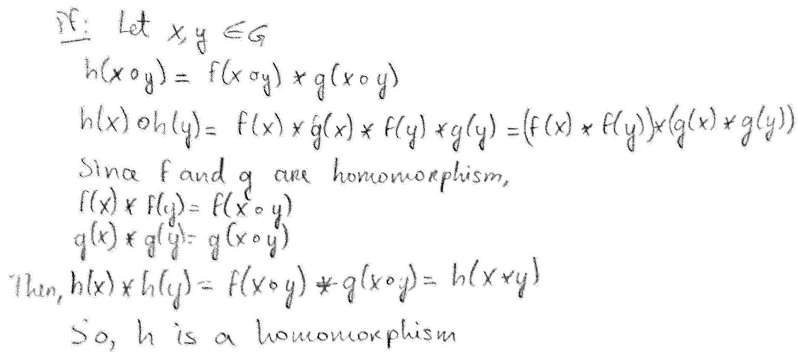


Figure 5. Student Response (Prompt 2, Version B: Not Abelian)

Discussion

Both the survey results and preliminary analysis of the case studies indicate that attending to the givens in a conjecture is a nontrivial task. A majority of undergraduate students unintentionally assumed a hypothesis was more general than given. The graduate student cases reflected similar (although slightly stronger) results.

Carissa is a case of hypothesis exploration that led to correcting her original proof error. More analysis of the graduate cases and follow-up interviews with the undergraduates could further flesh out this practice. Weber and Alcock (2004)'s syntactic and semantic reasoning could provide a framing for the reasoning occurring during this practice. If the generalizing/specializing arise from proof analysis, key idea (Raman, 2003) may also play an important role. In Carissa's case, a specific example was leveraged and so example usage in proving (Alcock and Inglis, 2008) may also provide an important lens to the specializing and generalizing practice.

Besides highlighting a common error, this study shines light on an often underdeveloped mathematical practice of generalizing and specializing. It is possible that instruction aimed at these practices may create better awareness of given statements and help minimize errors based on weakening theorems.

Questions for the Audience

1. What should be my next steps?
2. What frameworks may help inform the analysis of the practice of specializing/generalizing?

References

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics*, *69*(2), 111–129. doi:10.1007/s10649-008-9149-x

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, *38*(1), 85–109.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. *Research in collegiate mathematics education III*, *7*, 234–282.

Hart, E. W. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. *MAA NOTES*, 49–62.

Pólya, G. (1990). *Mathematics and plausible reasoning: Induction and analogy in mathematics* (Vol. 1). Princeton University Press. Retrieved from http://books.google.com.proxy.lib.pdx.edu/books?hl=en&lr=&id=-TWTcSa19jkC&oi=fnd&pg=PR5&dq=poyla++plausible+reasoning&ots=Ws7Z0g7a3B&sig=pPY8ZQJDzj6WKU1yzwf2Hggt-jk

Raman, M. (2003). Key Ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics*, *52*(3), 319–325. doi:10.1023/A:1024360204239

Selden, A., & Selden, J. (1987). Errors and misconceptions in college level theorem proving. In *Proceedings of the second international seminar on misconceptions and educational strategies in science and mathematics* (Vol. 3, pp. 457–470).

Selden, A., & Selden, J. (2008). Overcoming students' difficulties in learning to understand and construct proofs. *MAA Notes. Making the connection: Research and teaching in undergraduate mathematics education*, 95–110.

Simon, M. A. (1996). Beyond inductive and deductive reasoning: The search for a sense of knowing. *Educational Studies in Mathematics*, *30*(2), 197–210. doi:10.1007/BF00302630

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101–119. doi:10.1023/A:1015535614355

Weber, K., & Alcock, L. (2004). Semantic and Syntactic Proof Productions. *Educational Studies in Mathematics*, *56*(2/3), 209–234. doi:10.2307/4150282

Weber, K., & Mejia-Ramos, J. P. (2009). An alternative framework to evaluate proof productions: A reply to Alcock and Inglis. *The Journal of Mathematical Behavior*, *28*(4), 212–216. doi:10.1016/j.jmathb.2009.10.005

INSTRUCTORS' BELIEFS ON THE ROLE OF CALCULUS

Kathleen Melhuish
Portland State University

Estrella Johnson
Virginia Tech

Erin Glover
Portland State University

In this report we will draw on the Characteristics of Successful Programs in College Calculus data set in order to investigate instructor beliefs about the role calculus plays. Specifically, in this preliminary report, we have analyzed instructor interview transcripts in order to address the question: How do instructors perceive the role of calculus at successful four-year universities? Our preliminary analysis has uncovered six emerging themes. Each will be presented and illustrated with an instructor's quote.

Key words: Calculus, Teacher Beliefs, Instructional Practices

The Characteristics of Successful Programs in College Calculus (CSPCC) project is a large empirical study designed to investigate Calculus I. The primary focus of the CSPCC project is to identify factors that contribute to student success and understand how these factors are leveraged within highly successful programs. In addition to addressing these primary research goals, the CSPCC project has also collected much needed data about what happens in Calculus I across the nation. As discussed by Bressoud, Carlson, Mesa, and Rasmussen (2013), while Calculus I is offered at nearly every college and university across the nation, and taken by approximately 300,000 students every fall, very little data had been collected about what happens in Calculus I.

In this report we will draw on the CSPCC data set in order to investigate instructor beliefs about the role calculus plays within the university and their students' education. A great deal of literature has emerged connecting teacher beliefs about mathematics and their instructional practices (for instance: Cross, 2009; Thompson, 1992). However, much of this past research has largely considered teacher's beliefs on what constitutes mathematics and coming to know mathematics (as well as beliefs about students and epistemology in general). This study will consider a different set of teacher beliefs, those related to the purpose and role of a specific course. Specifically, in this preliminary report, we will be addressing the question: How do instructors perceive the role of calculus at successful four-year universities?

Background

The CSPCC project is a large empirical study investigating mainstream Calculus I in order to identify the factors that contribute to success and to understand how these factors are leveraged within highly successful programs. Phase 1 of CSPCC entailed large-scale surveys of a stratified random sample of college Calculus 1 classes across the United States. During Phase 2, explanatory case studies are being created. These cases were selected in part based in part on the results of the Phase 1 survey. Specifically, institutions were selected based on student *persistence* (continuing on to take Calculus 2), *success* (pass rates in Calculus 1), *and* reported increases in students' interest, confidence, and enjoyment of mathematics as a result of taking Calculus I. This second phase will lead to the development of a theoretical framework for understanding how to build a successful program in calculus and illustrative case studies for widespread dissemination. Eighteen institutions were selected as case study schools based on the results from the survey phase. The set of case study schools includes four community colleges, five bachelor's granting institutions, five master's granting institutions, and four PhD granting institutions. The research reported here is focused on the bachelors granting instructions.

Data and Methods

Our research team conducted site visits at five bachelors granting institutions: Urban State University (USU), University of Suburb (UoS), Private College (PC), Fake Catholic University (FCU), Regional University (RU). Four of the institutions were universities (two had very recently transitioned from colleges to universities). Three of these universities were private, while the fourth was a large urban public university. The final case study institution was a private liberal arts college. While on campus, we interviewed students, instructors, administrators, and others involved in the calculus program at the institution. This report will focus on the instructor interviews. We interviewed a total of 25 instructors over the course of the five case study site visits. Here, we will report on our ongoing analyses of the instructors' views about the role of calculus.

In order to address our first research questions, we analyzed two instructor interviews from each of the five universities. Initially, transcripts from the instructor interviews were read with the intent of identifying relevant excerpts. Most commonly these excerpts were found in response to interview questions regarding what the instructors like most/least about teaching calculus, what instructors want their students to get from their class, and what the instructors want students to get out assignments mathematically. From the 10 interviews 51 excerpts were identified. Open coding was then done of these 51 excerpts, resulting in six views of the role of calculus.

Preliminary Results

Here we will present the six emerging categories. After a brief description, we will provide illustrative quotes from instructors.

Calculus as a gateway - Calculus serving as a weed out course for either higher mathematics or client disciplines.

> ...it's the course that sort of tests their mettle and my mettle in terms of trying to teach them, have them learn the material and to be successful at the college level." (Smith, USU)

Calculus as a service course - Calculus serving to build skills needed in engineering and other client disciplines.

> I think there's interesting applications. So I think even though I'm a pure mathematician, I like to talk to the students about why they are sort of forced to be in this class. A lot of our students are not there by choice. And so well because there's applications to their majors, I think that's probably one of my favorite parts." (Bell, FCU)

Calculus as a tool for knowing mathematics - Calculus serving to introduce/strengthen mathematical practices.

> I also want them to appreciate the importance of being precise. That's part of also why I have them read the section of the book and make sure they understand, okay, hey, it's only five pages here maybe, but reading five pages could take 40 minutes if you really want to understand how we go from one step to the next. So mathematically understanding that all the statements that are made are made very carefully and each word means something." (Jones, USU)

Calculus for Calculus - Calculus as serving as an independently valuable subject.

> Well, for me, it's, you know, exciting ideas and being able to explain and transmit them, and that's especially true about Calculus I. I mean, I was just heading last night into the fundamental theorem of calculus. I was telling the students, 'This is one of the great ideas of Western civilization...' (Bianchi, UoS)

Calculus as the foundation for further mathematics - Calculus serving the role of building a mathematical foundation for more advanced courses.

> And then, after 2,000 years, Newton and Leibniz have this other idea that it's sort of a foundation of our pure and applied mathematics, and, you know, I find the ideas exciting, and I still do." (Bianchi, UoS)

6. *Calculus as the pinnacle of algebra* - Calculus serving as a capstone for the algebra sequence.

> So I think the algebra portion is a natural fit for kids, because most kids … regardless of their high school background usually have had a bunch of calculus – a bunch of algebra, so (inaudible) and algebra connects that, demystifying it. (Wells, PC)

Discussion and Plans for Future Research

Early analysis suggests that instructors view calculus as serving a multitude of roles. The pairs of instructors from each university often share similar views on the subject. For instance, both instructors from FPC emphasized the algebraic connections while this theme did not emerge in the other schools. Similarly, Calculus was mentioned as a foundation course by both instructors at the PC. Further analysis will be done with the remaining instructor interviews to further assess the correlation of role view amongst instructors at the same institution.

During the next round of analysis, we will consider possible relationships between an instructor's view on Calculus and their instructional practices and student assessments. Final exams have been collected from the various institutions and may capture ways in which instructor's views of the role of calculus could inform calculus instruction. Additionally, analysis on intended content to be covered in various calculus courses suggests that technology-focused schools may place more of an emphasis on applications (Johnson, Ellis & Rasmussen, 2014). This may suggest that instructors who view calculus as a service course may favor applications. Finally, because our preliminary analysis only considers four-year universities, instructor interviews from community colleges and PhD granting institutions could provide more insight into the role of the course. For instance, as transfer plays a much more significant role in community colleges, instructor views calculus may reflect this.

Questions for the Audience

1. How might instructors' views on calculus influence their teaching? Where should we look next?
2. Are these beliefs about different things (e.g., about calculus within mathematics vs. calculus within a students education)?

References

Bressoud, D., Carlson, M., Mesa, V., & Rasmussen, C. (2013). The calculus student: insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology*.

Bressoud, D., Carlson, M., Pearson, J., & Rasmussen, C. (2012). CSPCC4IJMEST-12-09-18.pdf. In *Preliminary Results of the Study of Characteristics of Successful Program in College Calculus*. Presented at the 12th International Congress on Mathematics Education., Seoul, Korea. Retrieved from http://www.maa.org/sites/default/files/pdf/cspcc/CSPCC4IJMEST-12-09-18.pdf

Cross, D. I. (2009). Alignment, cohesion, and change: Examining mathematics teachers' belief structures and their influence on instructional practices. *Journal of Mathematics Teacher Education*, *12*(5), 325–346. doi:10.1007/s10857-009-9120-5

Johnson, E., Ellis, J., & Rasmussen, C. (2014) How to make time: The relationships between concerns about coverage, material covered, instructional practices, and student success in college calculus. Seventeenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference. Denver, CO.

Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, *15*(2), 105–127.

Thompson, P. W., Castillo-Chavez, C., Culbertson, R., Flores, A., Greeley, R., Haag, S., Rutowski, R. L. (2007). Failing the Future: Problems of persistence and retention in science, technology, engineering, and mathematics (STEM) majors at Arizona State University. *Tempe, AZ. Office of the Provost*, *8*.

CHARACTERISTICS OF SUCCESSFUL PROGRAMS IN COLLEGE CALCULUS AT BACHELOR'S GRANTING UNIVERSITIES

Kathleen Melhuish, Sean Larsen, Erin Glover
Portland State University

Estrella Johnson
Virginia Tech

The CSPCC (Characteristics of Successful Programs in College Calculus) project is a large empirical study investigating mainstream Calculus 1 to identify the factors that contribute to success, to understand how these factors are leveraged within highly successful programs. Phase 1 of CSPCC entailed large-scale surveys of a stratified random sample of college Calculus 1 classes across the United States. From these surveys, successful institutions were selected as case studies. At each case study institution, Calculus I instructors, students and related administration were interviewed. In this report, we will present preliminary analysis on the five bachelor's granting institutions selected. We will discuss common themes and factors that have emerged from the five institutions.

Key Words: Calculus, Explanatory Case Study, STEM Student Retention

The MAA's project: CSPCC (Characteristics of Successful Programs in College Calculus) is a large empirical study aiming to identify key factors of successful calculus programs across the country. Calculus remains an essential course for nearly all STEM majors. With high rates of attrition in STEM majors (Lutzer, Maxwell, & Rodi, 2002), and a noticeable decline in students taking Calculus, it is vital that we consider successful Calculus models.

During the first phase of CSPCC, a large-scale survey was giving to a stratified random sample of mainstream Calculus 1 classes across the United States. Through analysis of these surveys (Bressoud, Carlson, Pearson, & Rasmussen, 2012), sixteen successful case study institutions were selected including four community colleges, four bachelor granting institutions, four masters granting institutions and four PhD granting institutions.

An initial pilot study was run (Larsen, Johnson, & Strand, 2013) identifying key features of a selected successful four-year institution. For this proposal, we will build on the pilot study highlighting commonalities found amongst the pilot institution as well as the four other selected institutions.

Methods

The five bachelor-granting university case studies were selected based on following measures of success that emerged during survey analysis. Regional University (RU), a private suburban university, was selected based on a low number of switchers, and high number of non-switchers as well as general high numbers on outcomes (persistence, grades, interest, confidence and enjoyment.) University of Suburb (UoS), a private suburban university, was selected based on outperforming expected pass rates by 12%, having a low number of non-switchers and high number of switchers and the highest outcome variables. Private College (PC), a private liberal arts college, was selected based on outperforming expected pass rates by 20%, and strong number of switchers/non-switchers. Urban State University (USU), a public urban university, was selected based on outperforming pass rate by 14%, above average outcomes and switcher/non-switchers and representing an urban public university. Fake Catholic University (FCU), a private urban university was selected based on positive attitude towards calculus during pilot and very high pass rates.

The CSPCC project team creating an interview protocol for case studies based on hypothesized areas of influence including: instructor attributes, departmental focus, classroom variables, and out of class expectations. At each institution, we interviewed calculus instructors (a total of 25 at our five case studies) and corresponding student focus groups, department chairs, placement coordinators, calculus coordinators, client disciplines, teaching learning center directors, college deans and other relevant personal totalling 66 interviews.

Through an initial analysis pass, each interview in the five case studies parsed based on excerpt relevance to institution (e.g. learning centers, department culture), classroom (e.g. assignments, technology), student (demographics, beliefs), instructor (attitudes, staffing), and outcomes (e.g. grades, persistence, etc.). For each case, two members of the research team highlighted relevant facts and features. A third member of team triangulated these documents to produce a summary of relevant features for each of the five cases.

A more complete analysis is now underway. Based on interview data and survey questions, a tagging scheme (24 tags) was developed amongst the CSPCC project team. Instructors, department chairs, calculus coordinators, and selected administrator interviews are tagged with tags such as assignments and assessments, outcomes, coordination, and student subjective characteristics. For example, if teacher said that, “My students often struggle with algebra, but the tutoring center provides good help for them,” the excerpt would be tagged *student subjective characteristics* and *learning resources.* Each interview is being tagged by two team members to assure no important topic is missed. The final tagging includes the union of both taggers. After, the tagging is completed, a more thorough analysis of cross-cutting features will be developed.

Through the coordination of initial highlighted facts and features for each university, data analysis based on relevant tags, exploratory case studies for each institution will be developed. The preliminary report will share emerging cross-cutting factors associated with the identified successful universities.

Preliminary Analysis

During early analysis, several cross-cutting features have begun to emerge.

Placement

Each of the five institutions emphasized proper placement of students and actively evaluated their placement policies. Three of the institutions utilized a placement test (USU - Accuplacer; FCU, RU - MAA placement test). The other two institutions did not use a placement exam, but did provide easy avenues for students to switch into appropriate classes. UoS gives a pretest at the beginning of the term and intentionally schedule their precalculus, calculus I and calculus II courses so students could easily switch into the appropriate class. PC utilized SAT scores and by policy created a very late drop-back date so students who are inappropriately placed could easily switch to the appropriate course.

Staffing

Most calculus courses were staffed by full-time faculty members with all schools having a large portion of sections taught by tenure-track faculty. At UoS, the course has historically been taught by 70% full-time faculty. Similarly, at RU, most calculus courses are taught by full-time faculty. At FCU and PC, a mix of full-time faculty and adjunct taught the course. AT USU, calculus courses are taught exclusively by full-time faculty.

Supporting Instructors

At all five institutions, instructor support was institutionalized. At FCU, they have a Teaching and Learning Collaborative to support, observe and provide feedback to other instructors. Faculty also receive generous funds for conferences. At RU, teachers are encouraged to be involved in MAA and they have a teaching center dedicated to helping teachers implement changes and assess results. At UoS, Faculty Center for Learning Development leads 12-20 workshops each term often emphasizing the use of technology to improve student learning. At USU, new faculty members are assigned mentors, there is a Center for Faculty Development which hosts discussion groups, learning communities and observes classes to provide instructor feedback. They also provide travel money for meetings and conferences. At PC, they have a Center for Teaching and Learning which runs many programs including teaching workshops and retreats, arranging meetings with new faculty cohorts for the first two years, and open Classroom project where faculty open up classroom for other faculty to observe.

Technology and Innovation

Our top four sites all fostered a culture of innovation with an emphasis on using technology. At UoS, the Faculty Center for Teaching and Learning assists teachers with the use of educational technology to promote learning. Currently the department is "flipping" half of the calculus sections under the funding of an NSF grant. At USU, technology use varied but instructors reported use of SMARTboards, clickers, WeBWork, projects, pencast videos, course wikis and Mathematica. At RU, a department committee was recently formed for the sole purpose of using technology in instruction. At FCU, the Teaching and Learning Collaborative provides support for instructional technologies. Technology varied by instructor, but several instructors reported the use of Geogebra in the calculus sequence.

Learning Centers

Learning resources were prevalent at each of the five institutions. All five colleges had on campus tutoring centers. At UoS, the Student Success Center matches students with tutors and academic support. The math department also has a tutoring center staffed by students. USU has a university-wide tutoring center staffed by students. FU has tutoring at the Academic Support Center for all subjects and the math department also has a drop-in tutoring center staffed by students. At FCU, the Learning Resource Center houses the Math Resource center which provides tutors for Calculus I students. PC's Center for Teaching and Learning houses a tutoring center where tutors work directly with Calculus I instructors.

Implications for Practice or Further Research

The goal of the CSPCC project is to create models of successful calculus with the hope of changing practice at institutions across the country. The explanatory case studies provide one form of these models. As we continue to analyze data, a more complete picture of successful calculus programs should emerge.

In the future, cross-cutting themes will be considered across various institution types with the hopes of isolating salient features of successful programs as well as considering the differences that may exist as a result of the nature of an institution.

Question for the Audience

1. What might be a good next step to continue analysis and build a complete case study model?
2. How could we go about analyzing the interviews in order to determine whether and how these characteristics support student success?

References

Bressoud, D., Carlson, M., Pearson, J., & Rasmussen, C. (2012). CSPCC4IJMEST-12-09-18.pdf. In *Preliminary Results of the Study of Caracteristics of Successful Program in College Calculus*. Presented at the 12th International Congress on Mathematics

Education., Seoul, Korea. Retrieved from http://www.maa.org/sites/default/files/pdf/cspcc/CSPCC4IJMEST-12-09-18.pdf

Larsen, S., Johnson, E., & Strand, S. (2013). Characteristics of Successful Programs in College Calculus: Pilot Case Study. Presented at the CRUME, Denver, CO. Retrieved from http://pzacad.pitzer.edu/~dbachman/RUME_XVI_Linked_Schedule/rume16_submission_68.pdf

Lutzer, D. J., Maxwell, J. W., & Rodi, S. B. (2002). Statistical abstract of undergraduate programs in the mathematical sciences in the United States. *Washington, DC: American Mathematical Society*. Retrieved from https://e-math2-ams-org.proxy.lib.pdx.edu/profession/data/cbms-survey/cbmssurvey-whole.pdf

CONCEPTIONS OF INVERSE TRIGONOMETRIC FUNCTIONS IN COMMUNITY COLLEGE LECTURES, TEXTBOOKS, AND STUDENT INTERVIEWS

Vilma Mesa
University of Michigan

Bradley Goldstein
University of Michigan

Abstract

We present a textbook analysis of conceptions of key ideas associated with inverse trigonometric functions using Balacheff's model of conceptions (Balacheff & Gaudin, 2010). We found conflicting conceptions of angles, trigonometric functions, and inverse trigonometric functions that may help explain difficulties that community college trigonometry instructors and their students face when explaining tasks associated with this topic. We make suggestions for further research.

Keywords: Trigonometry, Conceptions of mathematical notions, Curriculum

CONCEPTIONS OF INVERSE TRIGONOMETRIC FUNCTIONS IN COMMUNITY COLLEGE LECTURES, TEXTBOOKS, AND STUDENT INTERVIEWS

Trigonometry has traditionally been a high-school course, taught either as an independent course or as part of a pre-calculus course. Many post-secondary institutions offer trigonometry as part of a sequence of preparatory courses that lead to a calculus sequence (Lutzer, Rodi, Kirkman, & Maxwell, 2007). Over 55,000 students enrolled in a trigonometry course at two-year colleges in 2010 (Blair, Kirkman, & Maxwell, 2013, p. 137), yet we know very little about how this topic is taught. We have been aware of the paucity of research studies on trigonometric ideas, with very few scholars doing work on students' understanding of radian measure (Moore, 2010), discussing advantages and disadvantages of teaching a ratio or a functional approach to the trigonometric relationships (Kendal & Stacey, 1997; Weber, 2005), or describing future teachers' knowledge of trigonometry (Fi, 2003). As part of a larger project we studied how instructors explained inverse trigonometric functions to their students and also how students interpreted those explanations. Perhaps unsurprisingly, we found discrepancies between what teachers did in classroom and what their students revealed during interviews. In search for understanding this phenomenon, we decided to further investigate how textbooks present ideas associated with this topic. We sought to describe the conceptions of inverse trigonometric functions present in explanations in textbooks and by instructors and students, and use the discrepancies between these as a way to explain difficulties associated with these notions. We describe the theoretical framework that guided the investigation, the data collected and the analyses we performed, and our findings.

THEORETICAL FRAMEWORK

Balacheff's theory of mathematical conceptions (Balacheff & Gaudin, 2010) defines a *conception* as the interaction between the cognizant subject and the milieu (those features of the environment that relate to the knowledge at stake). His basic proposition is that conceptions of mathematical notions are tied to particular problems in which those conceptions emerge. Thus Newton's conception of function was substantially different than Dirichlet's because each was working with a different phenomenon (Balacheff, 1998). As mathematics develops and we solve new problems, our conceptions get transformed. The combination of all these different conceptions is what constitutes a persons' knowledge (*knowing*) about a particular mathematics notion. This way of understanding conceptions allows for potentially conflicting ideas about a mathematical notion to coexist without creating a dissonance in the knower. Indeed, discrepancies are only such for observers of the situation, as the knower might be using specific problems to conceptualize the mathematical ideas. Conceptions can be distinguished from each other because the problems (p) where they manifest require specific operations (r), systems of representations (l), and control structures (Σ, the organized set of criteria that helps the knower decide what to do in a given situation, determine that an answer has been found, and establish that the answer is correct decide whether an answer has been found and whether it is correct). This model of conceptions lends itself to an operationalization for analysis of different types of data.

DATA AND ANALYSIS

The data collection for this study took place between Fall 2010 and Summer 2011, in the context of three courses, trigonometry, pre-calculus, and calculus taught by two community college instructors, Elizabeth (trigonometry, pre-calculus) and Emmett (calculus). The teachers proposed the topic for investigation, inverse trigonometric functions, and agreed to (1) collect student data on knowledge of this topic prior to and after teaching a unit related to the topic in their courses, (2) let us interview their students after certain lessons had been delivered in order to collect information about their learning, and (3) discuss the findings with us, so we could together determine whether changes would be necessary. At the time of the data collection Elizabeth had seven years of college teaching experience, while Emmett had 16. We collected three types of instructional data: students' responses to a questionnaire on inverse trigonometric functions, interviews with students on various trigonometric ideas, and interviews with faculty on those results. We have video-recordings of classroom sessions and of the student and teacher interviews. In this study we focus on Elizabeth's presentation of inverse trigonometric functions, trigonometry and calculus students' interviews on a question addressing this topic. In addition we analyzed 10 textbooks: the trigonometry textbook that Elizabeth was using during the data collection period, and pre-calculus and calculus textbooks that could help us understand how notions related to inverse trigonometric functions were conveyed. As a group the textbooks represent a continuum of courses the students would take. We studied more than just the textbooks used at the college, because we wanted to have a broad perspective on how textbooks treated these topics. We performed the same analysis in each data source. First we identified portions of lectures, student interviews, and textbooks that related to inverse trigonometric functions. Then we identified each aspect of Balacheff's model of conceptions: the problems in which inverse functions are presented, the operations that are called for (e.g., Restrict the domain of sine, Define a new function, Sine, Switch the ordered pairs to obtain an inverse), the representations used (e.g., Unit circle, Triangle, Cartesian plane), and the control structure (e.g., Solution is in the correct interval, Identity is valid within the given range, Calculator outputs correct solutions). We repeated this analysis with the notions of angles and trigonometric relations in textbooks. These different analyses as gleaned from the textbooks help us make sense of Elizabeth's and the students' conceptions as revealed in lectures and responses to interview questions. The analyses allow us to discuss the imports of contradictory conceptions of inverse trigonometric functions in the textbooks and lectures, and how that permeates a student's conception of these notions.

FINDINGS

Teacher Explanation. Elizabeth begins her discussion about inverse trigonometry by discussing how (r) to "chop up" regular trigonometric functions to ensure that they are monotonic over the given interval. Initially, it seems that the domain of the sine curve [$-\pi/2$ to $\pi/2$] is the only part of the curve that gets translated to the inverse sine curve. Elizabeth uses a symbolic-graphical representation (l) of a unit circle, whose second and third quadrants are shaded in to emphasize that because the circle is shaded in on the left side, "we're not allowed to travel through quadrant two and three in order to get" (lines 199-200) to quadrant four. To illustrate what this means, Elizabeth brings an example that involves an angle outside of the graph of the function that had been restricted:

Elizabeth: So we're looking at the inverse cosine of -√2/2. So the answer here is not 5-fourths π, is it? It's outside the domain of inverse cosine. So we have to go up here and take the quadrant two angle that has the same output (lines 565-568).

Elizabeth explains that because we have "chopped off" part of the function, we cannot give a solution in that quadrant. While her solution is technically valid, she is implying that her solution is the *only* valid solution. She looses track that she "chopped up" the trigonometric functions to make them 1-to-1, *not* to restrict possible solutions. While Elizabeth usually discounts those areas outside of the domain of restricted sine, she sometimes uses expressions that reveal confusion about these "chopped off" sections. For example, Elizabeth explains that a calculator will only give output for the right hand side of the unit circle, "so anything happening with the sine that's happening in quadrant two or three, we have to compute ourselves using our brains" (131-132). Figure 1 illustrates how Elizabeth introduces the restrictions on sine to make it a monotonous function, and the answers she gives when confronted with an example outside the domain of restricted sine. Elizabeth is clearly aware of the tension between "chopped up" pieces as they are used to create a 1-to-1 function versus those same "chopped up" pieces when they arise in mathematical processes, but she does not explicitly articulate this tension to the class. The figure illustrates two competing conceptions that are related to the process of controlling that an answer has been found and that it makes sense.

Problem: *Interpret* sin^{-1}(*value*) *OR "for what angle* θ, *does sin*(θ)=*value?"*		
Conception 1		Conception 2
• Cartesian plane • Unit circle	**Representations**	• Triangle
• "Chop off" trig function • Shade regions of unit circle to determine which angles can be used • Use known facts of trig functions for specific angles	**Operations**	• Use Pythagorean theorem • Find missing lengths of sides of right triangle • Use ratio definition of trigonometric functions • Use the inverse function in the calculator
• Check that the value is within the range of the trigonometric function • Use brains • Check that the angle in the permitted quadrant • Inverse trig functions output angles.	**Control Strategies**	• Angles add up to 180 • Angles and values are nice and positive • Inverse trig functions output angles

Figure 1: Conceptions of inverse trig functions in Elizabeth's lecture.

Textbook Presentation. Most of the textbooks analyzed assume familiarity with angles. However, of the textbooks that discuss angles, each focus on *how to construct* or *measure* an angle (r) rather than on stating what an angle actually *is*. About half of the explanations give a dynamic view of angle using the arc length of a unit circle that is intersected by rays emanating from the origin of that circle (l). The other half of describe an angle in relation to the lines encompassing 'it' (l). The textbooks are similarly divided in their representation of regular trigonometric functions between two fundamentally different approaches. On one hand, trigonometric functions are seen as a direct translation from the *x*- and *y*-coordinates of the points at a given angle on the unit circle to the cosine and sine values, respectively, for those angle values (l). On the other hand, trigonometric relationships are represented as relationship between the sides of a right triangle (l). The unit circle method is presented as a more 'advanced method' than the triangle method and one that fits better the changing real world because of the way it incorporates periodicity. However, the translation from the (x,y) coordinate axes of the unit circle to the (x,y) axes of the sine graph is not well described. No textbook included an explanation for how angle values shift from an implicit polar system to become real numbers on the x-axis, or how the x-component of the point on the unit circle becomes a number on the y-axis (in the case of cosine). A problem (p) arises when constructing inverse trigonometric functions, because regular trigonometric functions are not increasingly monotonic on a given interval (they are not 1-to-1). When explaining why inverse sine has the range that is typically assigned to it ($[-\pi/2,\pi/2]$), textbook authors are generally divided between the two operations (r) used to obtain inverse sine. The first operation simply restricts sine on the domain from $-\pi/2$ to $\pi/2$. Reasons given for this restriction include, "this seems natural" or "this is agreed upon." A complication becomes apparent, however, when an example involving a function outside the range is presented. In that example, symbolic/algebraic representation (l) is used to bring the value into the accepted range. The second operation used to create inverse sine defines a new function, "Sine" to be sine restricted from $-\pi/2$ to $\pi/2$ (r). Inverse sine is not the inverse of the common sine function; rather, inverse sine is the inverse of the monotonic Sine function. This definition generates a new problem when solving an equation such as $Sin^{-1}x = y$ when y is any number outside the range of arcSine. None of the textbooks that had this definition of inverse sine addressed this type of problem. Figure 2 illustrates the connection between the opposing definitions of angles to similarly opposing definitions of trigonometry and ultimately, inverse trigonometric functions. In the figures, the parallel columns exhibit how a small difference in something as fundamental as angles can have significant consequences in the conception of inverse trigonometry, which is essentially built upon the idea of an angle.

Problem 1: *What is an angle?*		
Conception 1		Conception 2
• Line	**Representations**	• Rays
• Rotation around a point on the x-axis	**Operations**	• Intersection of two rays
• Not included	**Control Strategies**	• Not included
Problem 2: *What is sin of an angle?*		
• Unit circle	**Representations**	• Right triangle
• Identify coordinates of points on the unit circle. Sin of the angle is the y coordinate of the point.	**Operations**	• Identify the length of the sides of the triangle • Sin of the angle is the ratio of the opposite side to hypotenuse of the triangle
• The representation for sin(5pi/4) and sin(7pi/4) are different (another rotation) but they have the same value	**Control structures**	• The representation for sin(5pi/4) and sin(7pi/4) are the same (triangle is oriented differently)
Problem 3: *How is sin^{-1} defined?*		
• Unit circle • Cartesian plane	**Representations**	• Triangle • Cartesian plane

• Restrict sin to a monotonic interval in which it is monotonic	**Operations**	• Define a new function Sin-1 on the interval [-pi/2, pi/2]
• Other solutions exist because they are rotations on the unit circle	**Control**	• All solutions are encompassed here

Figure 1: Figure 2: Conceptions of angle, trigonometric and inverse trigonometric functions in the textbooks.

Student Interviews. In each of the student interviews, the students are presented with the trigonometric identities: $\cos^{-1}(\cos(x)) = x;\ 0 \le x \le \pi$ and $\cos(\cos^{-1}(x)) = x;\ -1 \le x \le 1$. The true reason for the bounds on *x* in the second identity is because those numbers bound the domain of arcsine (and the range of sine). Carl, a calculus student, however, draws a graph of cosine and explains:

Carl: Cosine between -π and π makes one complete cycle; makes one loop that touches -1 and 1 at both ends and –π and π but between -1 and 1 which is somewhere short of π over two on each axis, it still fails the horizontal line test so that is not one to one, so I do not know if that could be that couldn't be graphed I guess you could not create a function for that.

Carl is focused on the conditions needed for converting a function to its inverse function. He essentially gives a reason why $cos^{-1}(cos(x)) = x$ must be bound by 0 and π. He does recognize that when computing cosine of arccosine, there is no need to worry about monotonic functions anymore, because arccosine is defined to be the inverse of cosine, and thus defined to necessarily be one-to-one. Figure 3 shows Carl's circular reasoning about the nature of inverse trigonometric functions. He states that the cosine must pass the horizontal line test (be 1-to-1) in order to have a "proper inverse." When confronted with a value outside of the domain, he acts as if the restriction given is the *only* restriction possible and that a value outside of the domain would result in it "repeating itself over and over again" and cannot exist. He does not consider taking a different monotonous interval of the cosine function to find the value outside of the typical domain of inverse cosine.

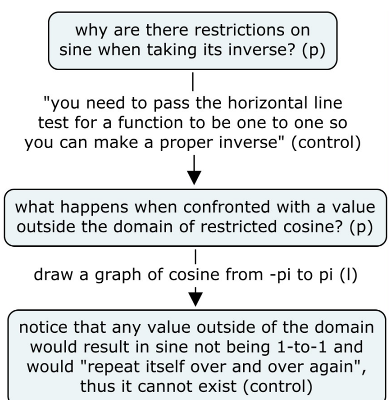


Figure 3: A calculus student conception of inverse trigonometry revealed in interviews.

DISCUSSION

This analysis of explanations given by a teacher, the textbooks, and a student highlights unresolved conflicts about notions of inverse trigonometric functions emerging from the very definition of angles. A typical trigonometric problem, find the angle of a given value, requires consistent definitions of angles, of trigonometric functions, and of their inverses. The control structures will vary depending on how these are defined and in some cases, it is problematic to account for all possible angles if one wants to recognize the periodicity of these functions. The crux is: Is there such a thing as an inverse sine? Do we need to define a new function? Is it a matter of convenience only? The ways in which the teacher argued for what needs to be done to find answers makes us think that these issues are not totally transparent. But they are not transparent in the textbooks either. We found different treatments to the issue, but in general we believe the presentations gloss over the complexity of reconciling these conceptions. The competing conceptions of inverse sine make it salient that these notions are problematic for students. Elizabeth's emphasis on "chopping" the function, an indirect reference to the need to restrict range and domain in defining the inverse masked the complex nature of what resolution was presented to the problem. The resolution does not seem to be completely satisfactory. Indeed, it is not clear what are the values of angles that one should be concerned about when finding answers to equations such as $\sin^{-1}(4_r)$. An intriguing result is the little attention that is given to the definition of angle in textbooks; this is an interesting finding in light of what Matos (1990, 1991) has identified as well, that angles have not been yet well defined. It is important to notice the impact of the definition on how trigonometric and inverse trigonometric functions are operationalized in textbooks and to explore the implications for teaching and learning.

LIMITATIONS AND AREAS FOR FURTHER RESEARCH

We acknowledge some limitations regarding the nature of the data we collected. It would have been useful to have the teachers comment on their explanations and to see whether they could identify the same discrepancies we identified. It would have been also useful to inquire about how they perceived the textbook's presentation regarding the nature of the inverse trigonometric functions. At the time of the data collection, we had other purposes in mind, which precluded us from collecting this information. A future study would include teachers' reflection on their own explanations and on the discrepancies in the book.

REFERENCES

Balacheff, N. (1998). Meaning: A property of the learner-milieu system. Grenoble.

Balacheff, N., & Gaudin, N. (2010). Modeling students' conceptions: The case of function. *Research in Collegiate Mathematics Education, 16*, 183-211.

Blair, R., Kirkman, E. E., & Maxwell, J. W. (2013). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States. Fall 2010 CBMS Survey*. Washington D.C.: American Mathematical Society.

Fi, C. (2003). Preservice secondary school mathematics teachers' knowledge of trigonometry: Subject matter content knowledge, pedagogical content knowledge, and envisioned pedagogy *Unpublished PhD dissertation*. Iowa City: University of Iowa.

Kendal, M., & Stacey, K. (1997). Teaching trigonometry. *Vinculum, 34*(1), 4-8.

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 CBMS Survey*. Washington, DC: American Mathematical Society.

Matos, J. (1990). The historical development of the concept of angle. *The Mathematics Educator, 1*(1), 4-11.

Matos, J. (1991). The historical development of the concept of angle , 2. *The Mathematics Educator, 2*(1), 18-24.

Moore, K. C. (2010). The role of quantitative reasoning in precalculus students learning central concepts of trigonometry *Unpublished PhD dissertation*. Tempe, AZ: Arizona State University.

Weber, K. (2005). Students' understanding of trigonometric functions. *Mathematics Education Research Journal, 17*(3), 94-115.

ACADEMIC AND SOCIAL INTEGRATION REVEALED IN *CHARACTERISTICS OF SUCCESSFUL PROGRAMS IN COLLEGE CALCULUS PROJECT*: THE TWO-YEAR COLLEGE CONTEXT

Vilma Mesa — University of Michigan
Nina White — University of Michigan
Helen Burn — Highline Community College

We present an analysis of features common across four Calculus I programs at two-year colleges identified as successful in the Characteristics of Successful Programs in College Calculus (CSPCC) study. In this paper we discuss how these features emerged in the analysis of the four cases and their connection to theories of student academic and social integration. Student academic and social integration have been identified as closely related to student persistence in college. We used a constant comparative analysis to identify themes within and across institutions, using transcripts of 22 interviews with faculty, staff, and administrators, and student focus groups. We discuss three of the seven major themes that arose, High quality instructors, Faculty autonomy and trust in the teaching of calculus, Supporting students academically and socially, and Attention to placement, which support a model of student academic and social integration. We present further research steps and some implications for practice.

Key words: Student persistence, Calculus, Two-year colleges

Given national need to strengthen our STEM-trained work force, understanding how institutions manage to keep students in the calculus track becomes an issue of national importance. The National Study of Calculus (Bressoud, Carlson, Mesa, & Rasmussen, 2013) collected survey data from over 200 institutions to describe program success, using measures at the student level regarding course standing (pass rates in Calculus I course and retention rates into Calculus II) and changes in student attitudes (growth in students' confidence, interest, and enjoyment in mathematics) over a semester of Calculus I. Based on these measures, the CSPCC team identified 17 successful institutions: five PhD granting, and four each of master's granting, baccalaureate granting, and associate's degree granting (see Hsu, Mesa, & The Calculus Case Collective, 2013 for details). The authors conducted intensive site visits at the two-year colleges in the sample (n=4) in order to identify features that contributed to being selected as a successful program and to understand these features in the context of the two-year college. Two-year colleges are particularly significant because they are responsible for a substantial portion of the undergraduate mathematical preparation, both in the form of remediation and in the preparation of the first two years of a STEM degree. Indeed, the latest figures report that 46% of students enrolled in a mathematics course in the U.S. are taking it a public two-year college (Blair, Kirkman, & Maxwell, 2013). This context is also significant because these institutions are increasingly seen as a viable pathway into STEM degrees. In this paper we sought to answer the following questions: What are common themes across the four two-year college cases that participants identify as being directly associated with the success of their calculus program? Do the themes suggest that successful programs help students integrate academically and socially into college? We present seven themes derived from our analysis and describe three in greater depth.

Theoretical and Analytical Frameworks

A number of scholars in higher education have hypothesized about the reasons students leave college without a degree. We use Tinto (1975) as the main model for our analytical framework; his work is derived from Durkheim's work on suicide (Durkheim, 1951). This model posits that an important feature of dropping out of college lies in the level to which students integrate academically and socially with their institutions: the more complex the network of relationships the individual develops with the institution, academically and socially, the less likely that the individual will drop out. Institutional features such as athletic teams, learning communities, and extracurricular activities are nowadays promoted under these premises. Research is mixed about the extent to which these models apply to non-residential campuses, such as two-year colleges, and whether these models can be used to interpret departure within specific academic programs, such as mathematics, and in particular course tracks, such as calculus.

A more nuanced conceptualization of the integration model suggests that integration is a process that evolves over time, and that it is possible to identify various phases of the process: separation, transition, and incorporation (Tinto, 1975, 1988, 1993). In early stages, students experience a disruption of ties and connections with their communities (exemplified by their moving from the high school to the college environment, or by leaving their hometown), followed by a period in which students start to understand the norms, behaviors, and rules of engagement in the new environment, which allows them to relativize the norms, behaviors, and rules they were used to, to then get into a period in which the students identify with, and own fully, the norms, behaviors, and rules of engagement of the new environment. Tinto posited that residential campuses had potential advantages in terms of this process, that the dual environment of home- or work-college could potentially compromise the full integration of the students. This could partially explain the lower graduation rates observed in two-year colleges. Further, the academic identity of two-year college students revealed in recent research (Cox, 2009a, 2009b) suggests the key role that faculty in two-year colleges play in both the academic integration and the social integration of students. In this paper we propose that some of the features of calculus programs that are common across the four successful two-year colleges in the CSPSS study offer corroboration of the importance of academic and social integration identified by Tinto, and that these features are especially relevant to two-year colleges in general.

Method

Data

The data used for the analysis reported herein come from interviews with key participants (instructors, department chairs, administrators) collected at the four two-year colleges included in the CSPCC sample. Table 1 shows institutional and departmental characteristics of the four colleges in the sample.

Table 1: Characteristics of the two-year colleges in our sample.

College	US Region	FTE[a] (2010)	FT:PT[b]	Number Calc I sections/term	Class Size
City College	Southeast	4,292	7:10	2	30-35
Urban College	Midwest	9,488	9:20	3 to 4	30
Rural College	West	2,788	7:0	1	30 (52[c])
Suburban College	Southeast	12,492	35:30	10	30 (15[d], 22[e])

a. Full-Time Equivalency student enrollment.
b. FT:PT is the ratio of Full-Time to Part-Time instructors in the mathematics department.
c. Number of students in the day of the observation, clearly exceeding the college's norm.

d. For honors sections.
e. Average by the end of the course.

Interviews were conducted with instructors teaching calculus, other instructors in the department (e.g., curriculum committee instructors), staff (e.g., people responsible for teaching or learning centers), and administrators (deans, chairs, etc.). The interview protocols were designed to understand elements affecting the Calculus I program from various levels: institutional (e.g. resources, communication with transfer universities or units), collegial (connections with other disciplines), departmental (e.g., coordination, selection of curriculum, assessment practices), and instructional (e.g., mode of instruction). We hypothesized that these features would allow us to understand the success measures (pass rates, retention rates, and changes in student enjoyment, confidence, and interest in mathematics) for each institution. The interviews lasted between 35 and 120 minutes and were transcribed. In addition, we observed as many Calculus I classes as possible, talked to groups of students in those classes, and collected classroom artifacts (syllabi, homework assignments, quizzes, and exams; see Table 2).

Table 2: Data Collected at the Four Community Colleges in the CSPCC Sample

College	Instructor Interviews	Other Interviews	Classroom Observations	Student Focus Groups
City College	3[a]	6	2	4 (43 students)
Urban College	5[b]	8	2	1 (26 students)
Rural College	1	5	1	1 (42 students)
Suburban College	8[c]	9	5	3 (39 students)

a. Includes interviews with a Calculus II instructor and a Calculus III instructor.
b. Includes two current Calculus I instructors and three instructors who have taught it regularly in the past.
c. Includes five current Calculus I instructors and three instructors who have taught it regularly in the past.

Analysis

The analysis of the data took place in two phases. The goal of the initial phase was to create a *Facts & Features* document that could be shared with our liaison at each college (typically the department chair) for member checking. The purpose of the *Facts & Features* document was to identify features that "key" interviewees described as contributing to the success of their calculus program. The 22 key interviews in this phase included the department chair and/or calculus-knowledgeable faculty (n=7), instructor interviews (n=9) and the majority of the interviews with high-level administrators (n=6) at each site. We limited the analysis to features that corresponded to actual practice (rather than opinions, judgment of those practices, or desired practices). Pairs of researchers generated summary lists of features for each of the interviews, which were then compared to create a consolidated list for each interview. Next, a master list of features was recorded for each institution, integrating each new interview analysis as it was completed. We used the student focus group transcriptions to corroborate and augment the features already identified. In the second phase, we analyzed the four *Facts & Features* documents to identify cross-cutting themes. To accomplish this, each author independently reviewed the documents and created broad categories such as "Faculty," "Students," and "Instruction/Content/Assignments" and distributed the features under each report into one of these categories. The large categories were further subdivided into more (and more) nuanced themes and new themes were created as needed (e.g., "Placement," "Learning Centers," "Transfer", "Interactive Lecture," "Diverse Student Characteristics," "Homework," "Student Preparation"). This was an iterative process. The themes that emerged were remarkably similar across researchers. Themes that arose in at least three of the four colleges form the basis of our findings. At this point, the analysis began to suggest that Tinto's model could be a useful lens to interpret some of the themes..

Findings

Our analysis generated seven themes that interviewees identified as related to the success of their calculus program and that were common across the colleges. We organized these themes into the three major categories shown in

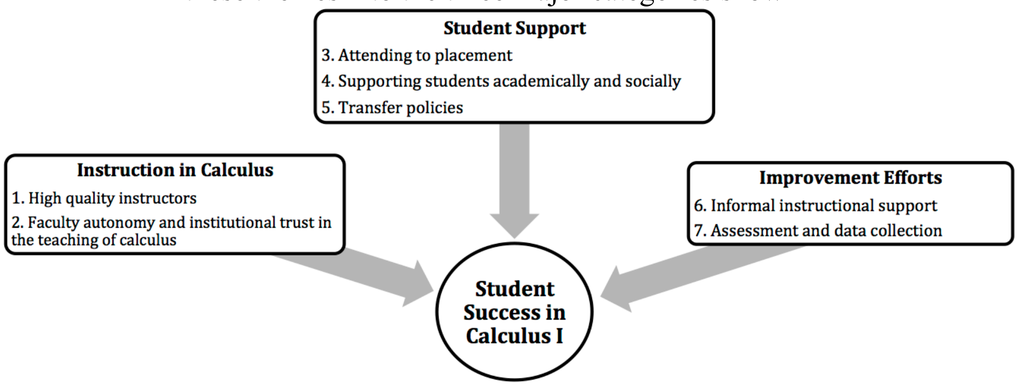


Figure 1: Instruction in Calculus, Student Support, and Improvement Efforts. While some of these themes can also be seen at other types of institutions, some were very specific to the two-year college setting, particularly themes 2, 5, and 6 (see Table 3).

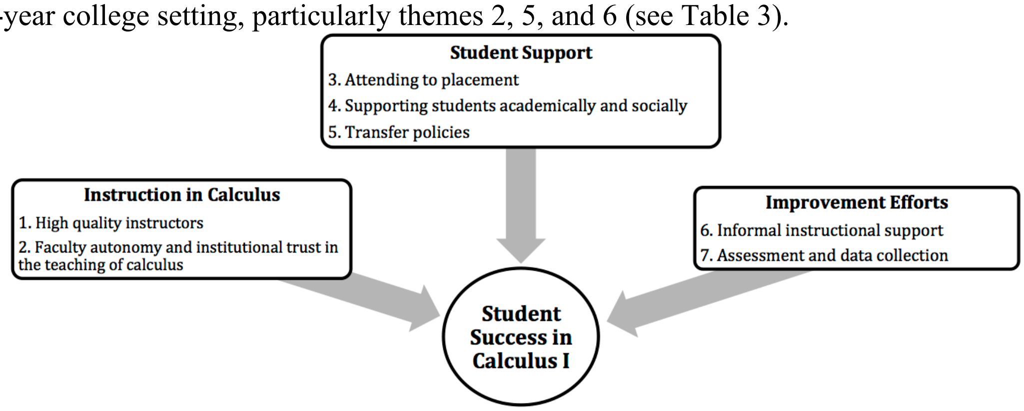


Figure 1: Major categories for themes identified in the cross-case analysis of two-year colleges.

Table 3: Cross-Case Analysis Themes.

1. High quality instructors	Cited as a major reason for program success. Descriptions of "quality" included: instructor availability and approachability, abundant content knowledge, and high expectations for developing conceptual understanding in addition to procedural competency/fluency.
2. Faculty autonomy and trust in the teaching of calculus	Instructors enjoyed much latitude and freedom in teaching calculus. Faculty and administrators trusted their colleagues to do the best for their students. No case had a department-wide policy about instructional approaches or technology. In cases with multiple calculus instructors, some amount of loose coordination across sections was sought through informal faculty collaboration as well as more formal calculus committees that provided loose coordination through common course outlines and/or common textbooks.
3. Attention to placement	All cases attended to effectively placing students into Calculus I. Three of the four sites had mandatory placement policies. Faculty were most satisfied with student preparation in the two cases where students tended to "course into" Calculus I by completing pre calculus.
4. Supporting students academically and socially	All cases supported students' academic success through learning centers, office hours, and advising. Socially, students had access to extracurricular math competitions and/or clubs as well as to study space. Instructors encouraged students to create calculus study groups.
5. Transfer policy	In all cases, the state played a role in ensuring course transfer to four-year institutions in the state or nearby states. Information about course transfer was conveyed formally through common course descriptions and articulation agreements (either between individual schools or coordinated at the state level) and informally through faculty members' personal knowledge and experience at four-year institutions.
6. Informal instructional support	In all cases, faculty reflected on their instruction individually and with their colleagues, mainly through casual conversation or email. This peer communication appeared to be the main mechanism for professional development and support. There were several examples of on-campus professional development opportunities and some support for faculty to participate in conferences, but they were not widely utilized by the mathematics faculty we interviewed.
7. Assessment and data collection	We observed three different types of data among our sites: (1) student learning outcomes at the college and departmental level, (2) student success in the college (e.g. pass rates, graduation rates), and (3) student success in transfer institutions. The colleges used the data mostly for reporting purposes. Faculty did not describe using the data for programmatic decisions.

Discussion

There exist theoretical perspectives that shed light on all of these themes. In this discussion, we use Tinto's model of student departure, which posits the importance of academic and social integration, as an analytic lens for understanding three of the themes in particular in greater depth. There is a high risk for academic and social *separation* for the two-year college population because the students live in two, sometimes, three worlds: work, family, and college. Tinto's model would suggest that this separation could facilitate drop-out, both at the institution and programmatic level. Three themes appear especially relevant in this context to mitigate this potential problem in separation.

High Quality Instructors

When our interview subjects used the word "quality" in describing their calculus instructors, their descriptions suggested that instructors were filling an important role in students' academic and social integration. For example, one of the oft-mentioned characteristics of the high-quality instructors teaching in these colleges was their commitment and dedication to the students. Instructors were described variously as approachable, available, and caring. Instructors likewise saw as their mission to assist their students the best they could. Rather than focusing solely on student learning gains or quality of curricular materials, most of the interview subjects (both student and faculty) described instructors' pedagogical strengths as providing support and encouragement to students. This seems to directly imply that instructors were filling an important role integrating students not just academically but socially as well, through their personal rapport and obvious dedication and caring.

Supporting Students Academically and Socially

Opportunities to receive academic support outside of class are key to students' academic and social integration in community college. Some forms of academic support we observed at our sites were institutional, such as campus-wide learning centers. More subtly, teachers and students often spoke of the importance of student study groups for best learning the difficult content of calculus. The importance of learning communities is well-known from Treisman's (1992) work with minority students, but also fits well into Tinto's model. In the colleges we visited, the importance of these groups was greater because in most cases the support available in the learning centers was limited to courses lower than calculus. In this context, the study groups became a social and academic support for the relatively smaller population of calculus students. All of our sites also had specific social supports within the context of the math department (e.g., math clubs, math teams, math competitions) that ostensibly support Tinto's model. However, it was unclear the extent to which students participated in these offerings, so they are not relevant to our main point. This is in contrast to informal study groups described above, in which students in focus groups readily reported participating actively.

Attention to Placement

The first step in students' academic integration in community colleges is through course placement. All cases attended to effectively placing students into Calculus I. Specifically, at the two smaller sites, department chairs reported that virtually all of their Calculus I students took Precalculus at the same institution, rather than placing directly into Calculus I via a placement test. This "in-house" preparation was described by the Calculus I instructors as very effective. While there are certainly curricular forces at work in these cases, Tinto's model also helps explain why this might be effective; students are going through a program *together* with a common experience, and have become a community of learners by the time they reach Calculus I.

Implications and Future Directions

This paper focused specifically on how the process of integration described in Tinto's (1975, 1988) model was supported in the successful sites we visited. The importance of this paper lies in part in exposing the mathematics education community to Tinto's model, which has aspects of the "emergent perspectives" by combining cognitive and social elements of learning (Cobb & Yackel, 1996) but extending its reach to the program level. The findings have implications for practice for mathematics faculty teaching calculus at two-year colleges in helping them consider the extent to which they collectively attend to students' social integration either within the classroom or through out-of-class support. Encouraging participation in informal study groups can be a simple strategy that can be tried out. Departments or individual faculty can use the examples of social and academic integration that emerged from our themes as a starting point for such an exploration.

This line of study has clear future directions in both breadth and depth. In terms of depth, augmenting Tinto's model with related theories from higher education, such as "involvement" (Astin, 1984), "engagement," (Kuh, 2008) and "validation" (Rendon, 2006) could enhance our theoretical framework in future analyses.

In terms of breadth, there were observed themes (e.g., transfer policies or use of data) that may require different theoretical perspectives to explain their prominence. Incorporating more areas of theory will allow us to better understand the remaining four of the seven identified themes. For example, the interest these institutions had in collecting data resonates well with the notion of "continuous planning" for improvement described by Briggs and colleagues (2003). Yet it is unclear that there was clarity within institutions about the value of the data collected.

Also, in addition to having the "program" as the unit of analysis, we plan to consider the "department" and the "institution" and their roles in facilitating student success. The

CSPCC interview data with department and college administrators could help describe the ways in which not just key faculty but other institutional players contribute to students' academic and social integration. The theoretical perspectives helpful in understanding a "successful" program (e.g. Tinto) would apply equally for these contexts.

References

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel, 25*, 297-308.

Blair, R., Kirkman, E. E., & Maxwell, J. W. (2013). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States. Fall 2010 CBMS Survey*. Washington D.C.: American Mathematical Society.

Bressoud, D. M., Carlson, M., Mesa, V., & Rasmussen, C. L. (2013). The calculus student: Insights from the MAA national study. *International Journal of Mathematical Education in Science and Technology*. doi: 10.1080/0020739X.2013.798874

Briggs, C. L., Stark, J., & Rowland-Poplawski, J. (2003). How do we know a "contiuous planning" academic program when we see one? *Journal of Higher Education, 74*(4), 361-385.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*, 175-190.

Cox, R. D. (2009a). *The college fear factor: How students and professors misunderstand one another*. Cambridge, MA: Harvard University Press.

Cox, R. D. (2009b). "It was just that I was afraid": Promoting success by addressing students' fear of failure. *Community College Review, 37*, 52-80.

Durkheim, E. (1951). Egotistic suicide (J. A. Spaulding & G. Simpson, Trans.). In E. Durkheim (Ed.), *Suicide* (pp. 141-151). Glencoe: Free Press.

Hsu, E., Mesa, V., & The Calculus Case Collective. (2013). *Synthesizing measures of institutional success*. CSPCC-Technical Report #1. Mathematical Association of America. Washington D.C.

Kuh, G. D. (2008). *High impact eduational practices: What they are, who has access to them, and why they matter*. Washington, D.C.: American Association of Colleges & Universities.

Rendon, L. (2006). *Reconceptualizing success for underserved studens in higher education*. Response paper for NPEC National Post-secondary Education Cooperative. NPEC National Post-secondary Education Cooperative. Retrieved from http://nces.ed.gov/npec/pdf/resp_Rendon.pdf

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89-125.

Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of student leaving. *The Journal of Higher Education, 59*(4), 438-455.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Treisman, P. U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *College Mathematics Journal, 23*, 362-372.

100

FACTORS ASSOCIATED WITH THE SUCCESS OF FEMALE MATHEMATICS DOCTORAL STUDENTS

Emily Miller
University of Delaware

Although the gender gap in participation in undergraduate mathematics has narrowed, disparities persist at the doctoral level. Only 30 percent of recent doctoral recipients in mathematics were women (Hill, Corbett, & St. Rose, 2010). To increase retention of women in mathematics doctoral programs, it is critical to study the factors that are associated with success. A survey was distributed to 142 female mathematics professors to assess the impact of factors that could have contributed to their success. Results point to changeable factors that can be implemented to narrow the gender gap. Salient factors include persistence and dedication, strong undergraduate preparation and quality doctoral courses, and support from the doctoral advisor. Results show that gender still has an impact on the experiences of the participants, but there may be reason for optimism. Respondents who received their doctorates more recently reported less gender discrimination.

Keywords: Gender equity; doctoral mathematics; retention

Although the gender gap in participation in undergraduate mathematics has narrowed, disparities still exist at the doctoral level. Only 30 percent of recent doctoral recipients in mathematics were women (Hill, Corbett, & St. Rose, 2010). Women attempting to obtain doctorate degrees in mathematics complete degrees at a lower rate than men (Berg & Ferber, 1983), and even those who do complete their Ph.D.'s take longer to do so, on average, than men (Herzig, 2004). By identifying factors that affect the quality of women's experiences, doctoral programs can be improved to increase the success of this group.

Prior research has focused on attrition. Given that retention can be viewed as the absence of attrition, it is reasonable to hypothesize that factors, operating in a reverse direction, might help retain female doctoral students in mathematics. There have been few studies that examine this hypothesis. In this study, I identify factors that affect retention, as reported by female mathematics professors, using a large-sample quantitative survey methodology.

Theoretical Framework

Although there is no model of doctoral persistence specific to mathematics, Tinto (1993) proposes a three-stage framework to examine doctoral persistence across all fields. The first stage, transition and adjustment, lasts for about the first year of doctoral study and is the process of socialization and accommodation into the norms of the graduate department. The second stage, development of competence, usually begins at the second year of the doctoral program and ends as the student passes her candidacy examinations. The initiation of the third stage signifies that the student has acquired the knowledge and skills required to conduct doctoral-level research, in the form of a dissertation. The third stage, dissertation completion, culminates with a successful dissertation defense.

A review of the relevant literature led to a list of 16 factors that could contribute to attrition. Two additional factors were added based on my own hypotheses. These 18 factors can be organized based on theoretical framework, and partitioned into two groups: (a) personal factors, which are unique to each student, and (b) program factors, which describe characteristics of the student's doctoral program. Each of these factors has the potential to affect the student's ability to be successful in her doctoral program in mathematics.

Goals and Research Questions

A desired result of this research is to identify factors that women view as critical to their success in obtaining a Ph.D. in mathematics so that these factors can be used to inform the design of doctoral programs that facilitate women's success. My research questions are: (1) What are the experiences of successful female mathematics doctoral students, as reported by female mathematics professors, and how do they describe the effects of those experiences? (2) What factors are most important to the success of female mathematics doctoral students, as reported by female mathematics professors?

Methods

Sample

The desired population for this study consists of female, tenure-track mathematics professors with doctorate degrees in mathematics who are employed in prestigious universities in the U.S. To obtain a sample from this population, the top 85 universities were selected from *U.S. News & World Report*'s 2012 National University Rankings (U.S. News & World Report, 2012). The mathematics faculty listings for these top universities were then obtained online. Email addresses of potential participants were identified from these listings.

Although this sample is not randomly selected, it is sufficient for this study. In searching for factors that allow women to be successful in doctoral mathematics programs, it is relevant to study those who were particularly successful, as evidenced by their hiring in tenure-track positions at prestigious universities. If, despite their success, the women in this sample still identify aspects of their doctoral program that hindered their progress, this may be even more meaningful in understanding the factors that affect retention.

Participants

Survey invitations were sent electronically to 392 potential participants. Ten invitations were undeliverable because of inaccurate email addresses. Of the remaining 382 invitations, 182 responses were received, for a response rate of 47.6 percent. Thirty-one responses were omitted from data analysis because the respondents did not fit the selection criteria. Nine respondents did not complete the survey and were not included in data analyses. This left a sample size of 142 available for data analysis. Forty one percent of the respondents are full professors, 24 percent are associate professors, and 35 percent are assistant professors. Ninety-two percent of the respondents work at institutions that grant Ph.D.'s in mathematics.

Instrument Design

A survey instrument was designed based on findings from a literature review. The survey has four blocks of questions. The first block gathered demographic information and was included to ensure that only the appropriate participants completed the survey. The second block consisted of open-ended questions, asking about what the participants believe made them successful and what challenges they overcame to receive their doctorate. Two open-ended questions were also included at the end of the survey. The first asked about the respondents' teaching style preferences, and the second asked for any additional thoughts.

The third block of questions utilized a two-step format. First, the respondents were presented with 18 factors that may have contributed to their success. The participants were asked to indicate whether they experienced these factors during their doctoral program by answering "True" or "False" for each statement. If the respondents selected "True," they were directed to a Likert scale, ranging from "Extremely negative effect on my ability to be successful" to "Extremely positive effect on my ability to be successful." If the respondent selected "False," they were directed to a separate five-point Likert scale, with identical labels. This scale asked respondents to assess the impact on their success of the factors they stated were not present during their doctoral program. Each of the factors was now reworded in the negative to more accurately reflect the respondent's experience. Finally, in the fourth block, the respondents were presented with the factors and were asked to choose the five factors that were most influential in their success and to rank them in order of importance.

Results

RQ1: The Experience of Female Mathematics Doctoral Students

The personal factor of persistence and dedication, along with the program factors of strong undergraduate content preparation and high quality graduate instruction were the most impactful on the experiences of the respondents in this sample. Comparing the "True" and "False" lists of personal factors that were rated significantly higher or lower than the others, the impact of persistence and dedication is apparent. This factor appears as having one of the highest ratings (indicating a positive effect on the respondents' ability to be successful) for those who had it, and one of the lowest negative ratings (indicating a negative effect on the respondents' ability to be successful) for those who did not. Similarly, for the program factors, the quality of the content preparation from the respondents' undergraduate program and the quality of the courses the respondents took while enrolled in their doctoral program appeared on both lists. Those who had high quality instructional experiences at the undergraduate and graduate level found it to have a highly positive impact on their success, and those who did not found it very detrimental.

RQ2: Importance of Factors to the Success of Female Mathematics Doctoral Students

The statement "I was persistent and dedicated to my work during my doctoral study in mathematics" was the highest ranked personal factor in terms of importance, with a mean of 2.61. For the program factors, the highest ranked was "I received strong support and encouragement from my doctoral advisor in mathematics," with a mean of 2.24. The program factor "I received strong support and encouragement from my doctoral advisor in mathematics" was ranked first most often overall, with 30 out of 134 respondents (22.4 percent) selecting that factor as most important to their success. The personal factor "I was persistent and dedicated to my work during my doctoral study in mathematics" was ranked as most important by 26 out of 134 respondents (19.4 percent), second most often overall and the most often out of the six personal factors. The personal factor, "I was persistent and dedicated to my work during my doctoral study in mathematics," was the most common factor selected for the top five for the ranking question, with 99 out of 134 respondents (73.9 percent) ranking that factor.

The Impact of the Doctoral Advisor

When asked about the most important contributors to their success in obtaining a doctorate in mathematics, nearly half of respondents (46.7 percent) included the support of their doctoral advisor in their ranked list. Furthermore, the support provided by the doctoral advisor was second only to persistence in the mean ranking of importance. Similar results can be seen from the Likert data, where nearly all of the respondents who agreed that their advisor was supportive said that this had a positive or extremely positive effect on their ability to be successful in their doctoral program. The open-ended data was used to gain a more comprehensive description of the participants' meanings of support from their doctoral advisor. Specifically, they spoke about receiving mentoring, assistance in choosing an appropriate thesis or dissertation topic, helpful and timely feedback on work, recognition of progress through the program, and financial support through grants and fellowships.

However, an overwhelming majority of respondents (95.8 percent) stated that their doctoral advisor was male. Only 6 respondents (4.2 percent) had a female doctoral advisor. This is most likely due to the dearth of female professors in most university math departments, evidenced by the fact that 85.2 percent of respondents expressed that they did not have enough female mentors available to them. Over a third of the sample responded that this lack of female mentors had a negative effect on their ability to be successful.

The Impact of Undergraduate Preparation

Undergraduate mathematics preparation was posited to be a factor that could have a large effect on the success of women in doctoral programs. Almost a third of the sample claimed they did not have strong content preparation in mathematics from their undergraduate program. Eighty-two percent of these women reported this factor had an extremely negative or negative effect on their ability to be successful in their program. Sixty nine percent reported having strong content preparation in mathematics from their undergraduate program, and 94 percent of this group reported this factor had an extremely positive or positive effect on their ability to be successful in their doctoral studies. Whereas 90 percent of the sample attended a university in the United States during their time as a doctoral student, 48 percent of the women in the sample received the entirety of their undergraduate training in mathematics in countries other than the United States. Women who received their undergraduate training in mathematics in the United States were more likely to report that they did not receive strong content preparation at the undergraduate level. Forty out of 74 women (54.1 percent) who had undergraduate preparation in the United States claimed that their preparation was strong, while 56 out of 65 women (86.2 percent) who were prepared for their doctoral studies outside of the United States reported strong preparation. Results of a 2 x 2 chi-squared analysis show a statistically significant difference between groups (chi-squared = 16.689, $df = 1$, $p < .001$). In fact, the odds ratio for this data is 5.31, meaning that women who had undergraduate preparation in the United States were over five times as likely to report weak preparation than those who did not.

Reports of Gender Bias

Even though the number of women choosing to pursue mathematics has been rising, gender continues to exert an influence on their experiences in mathematics doctoral programs. Approximately 17 percent of respondents believed the other students in their program thought they were less likely to be successful than a male student, while a similar percentage of respondents (16 percent) had the same belief about the professors in their doctoral program. About three-quarters of respondents holding each of these beliefs felt that the inequity had a negative effect on them. This suggests negative stereotypical beliefs about the competency of women in mathematics still pervade some women's experiences.

Although they were not prompted to discuss instances of gender bias or discrimination, many respondents nevertheless decided to write about this topic in their open-ended responses. Twenty-five respondents (17.6 percent) described an occurrence of overt gender bias, whereas 20 respondents (14.1 percent) explicitly stated they had not experienced gender bias or made similar gender-blind statements. The instances of gender bias were summed up by one respondent as "a nagging sense that [she] was intruding on a 'boy's club.'" However, these experiences were not the case for everyone. Almost as many respondents stated in their open-ended responses that they had never experienced any form of bias in their doctoral program. One additional unexpected finding also arose from the open-ended responses. Although not prompted to discuss this on the survey, six participants described their belief that the gender discrimination they experienced worsened as they transitioned into an academic career. One respondent stated: "The discrimination just gets worse and worse the higher you go." Although this may not be surprising, given that the numbers of women dwindle even further at the faculty level, this may be an area for further study, given the number of women that volunteered this comparison without being directly asked.

Discussion

Although women have made great strides in equalizing the gender gap in participation in mathematics at the undergraduate level, the environment of doctoral level mathematics is still dominated by men. In order to increase the numbers of women pursuing advanced degrees in mathematics, it is necessary to evaluate the personal and program factors that help to retain women in these programs.

In this study, a quantitative survey methodology was used to assess the impact of 18 factors on the success of female mathematics doctoral students. The sample consisted of female mathematics professors employed at prestigious universities in the United States. The impact of persistence and dedication was found across data sources, as well as the importance of strong undergraduate preparation and quality courses at the graduate level. These three factors were rated as highly beneficial for those participants who experienced them, and extremely detrimental for those who did not. However, in a departure from previously published results, there was little agreement on what the participants considered to be good teaching. The responses were wide-ranging, suggesting that efforts to retain women should focus on other aspects of mathematics doctoral programs. The importance of having a supportive doctoral advisor also appeared as an influential factor.

The importance of strong undergraduate preparation arose as one of the most striking findings of this study, since women who received their undergraduate degree in the United States were much more likely to report that their undergraduate training in mathematics was weak, when compared to those who received their undergraduate degree elsewhere. This has implications for how American universities are preparing their students for the transition to graduate study. If female students from the United States are to be successful in doctoral study in mathematics, their preparation must become on par with their international counterparts.

Although personal characteristics are outside the scope of the doctoral program's influence, improvements made to program factors can ameliorate the impact of any negative personal factors. In particular, confidence, motivation, persistence, and dedication could be improved through the implementation of student-student or student-faculty mentoring programs. This type of socialization could be provided through university chapters of the AWM; however, only 19.7 percent of respondents reported the availability of such a program during their doctoral study. Issues related to family responsibilities can be improved through greater understanding and flexibility on the part of the graduate department. Since the age of graduate study for most women coincides with young adulthood, when many choose to start families, an essential part of making the mathematics doctoral program environment welcoming for women is becoming more accepting of the competing "greedy institutions" of academic and family life (Herzig, 2010, p. 198).

Although Tinto's (1993) longitudinal model of graduate persistence was not designed to be specific to female students in male-dominated disciplines, the alignment of the model with the results of this is apparent. In the first stage, transition and adjustment, the female mathematics doctoral student must determine the compatibility of her expectations with her experiences in the graduate department. Early instances of gender bias may cause her to deem the environment unsuitable. Even if she stays, these experiences could decrease her persistence or confidence, or commitment to doctoral study. During the second stage, development of competence, her qualifications as a mathematician are called to the forefront. This could serve to reawaken questions of inadequate undergraduate preparation. This stage is also focused on socialization and integration, which may be more difficult for female students with a lack of female mentors or role models with the mathematics department. Lastly, in the final stage, dissertation completion, the role of the advisor becomes central to the student's experience. With a supportive advisor, the student may very well thrive and complete her degree. However, with an unsupportive or unsuitable advisor, the student may leave her program at this advanced stage.

References

Berg, H.M., & Ferber, M.A. (1983). Men and women graduate students: Who succeeds and why? *The Journal of Higher Education*, *54*(6), 629-648.

Herzig, A.H. (2004). Becoming mathematicians: Women and students of color choosing and leaving doctoral mathematics. *Review of Educational Research*, *74*(2), 171–214.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. American Association of University Women.

National University Rankings. (2012). *U.S. News & World Report*. Retrieved August 24, 2012, from http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-universities.

Tinto, V. (1993). Toward a theory of doctoral persistence. *Leaving college: Rethinking the causes and cures of student attrition* (pp. 230–243). Chicago: University of Chicago Press.

101

TEACHING THE CONCEPT OF MATHEMATICAL DEFINTION USING STUDENT CONSTRUCTION AND SELF-ASSESSMENT

Susanna Molitoris Miller
Kennesaw State University

Definitions are an important part of the study of mathematics, yet many students struggle with successfully understanding and using this construct. It has been suggested that students may improve their understanding of mathematical definitions by engaging in the act of writing definitions (de Villiers, Govender, & Patterson, 2009)*. Through a mixture of survey and teaching experiment methodology this study explores pre-service elementary teachers' understanding of mathematical definitions before and after engaging in a teaching experiment which provided many opportunities for the participants to write their own mathematical definitions for familiar and novel classes of quadrilaterals. Definitions were assessed as having necessary, sufficient and minimal conditions. It was found that while many students initially struggled to write definitions that meet these qualifications, the process of trying to construct their own definitions did improve students' understanding of these characteristics of mathematical definitions.*

Key words: Definitions, Geometry, Teaching Experiments, Elementary Teacher Training

Mathematical definitions play an important role in the study of practically every area of mathematics (de Villiers, Govender, & Patterson, 2009; Usiskin & Griffin, 2008; Vinner, 1991). They serve as a means of communicating about mathematical ideas by describing concepts in a precise and efficient manner which would allow the reader to differentiate the concept being defined from other similar concepts (de Villiers, Govender & Patterson, 2009). In axiomatic systems, the way a concept is defined can affect which statements are regarded as given facts and which must be proved as theorems and corollaries (Usiskin & Griffin, 2008).

Difficulties with Mathematical Definitions

Unfortunately, many students struggle with the important concept of mathematical definition. Zaslavsky and Shir (2005) found that 12^{th}-grade students have mixed and varied conceptions of the form and purpose of a mathematical definition. The students in their study questioned the importance of minimal sets of conditions, whether or not any statement containing necessary and sufficient conditions could serve as a definition, which properties of an object can be included in a definition, and if there can be more than one correct definition of a concept (Zaslavsky & Shir, 2005). Several studies have shown that during their training pre-service teachers also have a deficient understanding of the form and purposes of mathematical definitions (e. g., Leikin & Winicki-Landman, 2001; Linchevsky, Vinner, & Karsenty, 1992). Fujita (2012) found that pre-service elementary teachers were not fluent with the concept of mathematical definition as evidenced by their reliance on prototypical examples rather than applying definitions when solving problems. Chesler (2012) found that pre-service secondary teachers near the end of their preparation program struggled to complete tasks involving selecting, applying, comparing, and interpreting definitions of a function.

Research Question

In response to these difficulties, there have been some suggestions as to how we might better assist students in understanding and using mathematical definitions. De Villiers, Govender and Patterson (2009) encouraged allowing students to engage in the process of constructing mathematical definitions, conjecturing that such experiences could improve students' understanding of "geometric definitions, and of the concepts to which they related" as well as " the nature of definitions" (p. 201). I designed a qualitative study involving a blend of survey, teaching experiment, and case study methodologies to explore the potential of that recommendation. This paper/presentation will address the following research question: How do students' understandings of mathematical definitions change after they engage in activities involving the composition of mathematical definitions?

Characteristics of a Mathematical Definition

Throughout this study I use the term "mathematical definition" to refer to definitions with the following characteristics as described by de Villiers, Govender, and Patterson (2009). These authors state that "definitions are very concise, contain technical terms, and require an immediate synthesis into a sound concept image" and that definitions function as "tools for communication, for reorganizing old knowledge, and for building new knowledge through proof" (p. 189). They recognize definitions as "human 'inventions'" (p. 191) which can come in a variety of types. One distinction concerns how definitions relate to the relationships between sets. A *hierarchical* (or inclusive) definition "allows the inclusion of more particular concepts as subsets of the more general concept" but *partitional* (or exclusive) definitions describe a system in which "the concepts involved are disjoint from each other" (p. 191). They consider a definition "correct" if it contains both *necessary* and *sufficient* conditions. These attributes mean that "the condition applies to all elements of the set" and "whenever [the condition] is met, we obtain all the elements of the set we want to describe" respectively (p. 193). Definitions can be "economical" if they contain "a minimal set of necessary and sufficient conditions" but "uneconomical" definitions contain "redundant properties" (de Villiers, Govender, & Patterson, 2009, p. 196). The following study contributes to the filed by exploring pre-service teachers' understanding of the structure of mathematical definitions focusing specifically on necessary, sufficient, and minimal sets of conditions.

Methods

I designed this study using a combination of survey and teaching experiment methodologies in order to explore pre-service teachers' understanding of mathematical definitions via their abilities to write mathematical definitions for various types of quadrilaterals. In the first part of the study I used survey data from 44 elementary education majors to get an overview of how pre-service teachers think of mathematical definitions by exploring how they define quadrilateral types before formally studying either definitions or geometric shapes on the college level. In the second part of the study I performed a conjecture driven teaching experiment (Confrey & Lachance, 2000) with three students who were carefully selected as case studies based on the results of the initial survey. The teaching sessions focused on activities which relate a definition to the set of figures to which it corresponds as a means of improving students understanding of the structure of a mathematical definition including necessary sufficient and minimal conditions. A post-measure survey was also used to assess student growth or changes.

The teaching experiment portion of this study was based on a learning trajectory informed by three theories, shown in Figure 1. The first theory, the Conceptual Learning and Development Model, is a theory describing stages for the acquisition of concepts in general (Sowder, 1980). The second theory, the van Hiele theory of development of geometric reasoning, is a theory which also includes stages but is focuses specifically within the context of geometry (Burger & Shaughnessy, 1986). The third theory, concept image, applies to the learning of concepts in general and describes the non-linear accumulation and organization of knowledge over time (Vinner, 1991). Multiple theories were used in order to explore multiple facets of how students learn about mathematical definitions. Previous work has often focused on only one theory related to one component of the learning trajectory while this study highlights the importance of all three sets of arrows included in the trajectory as well as the

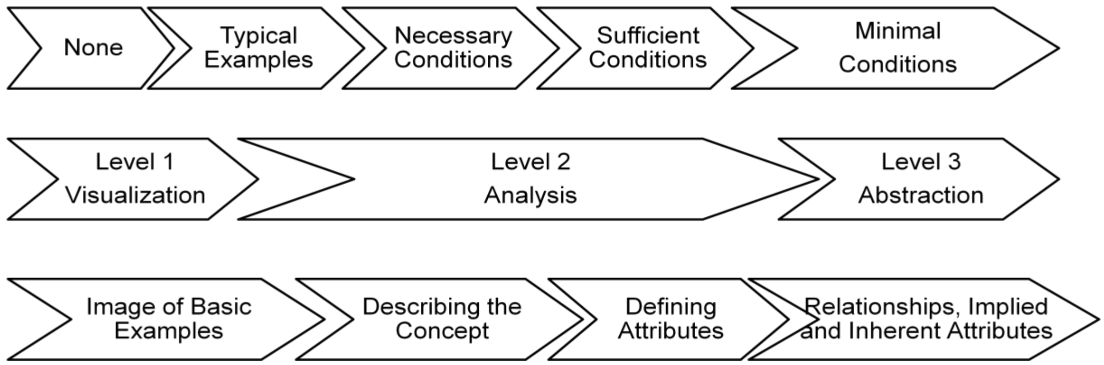


Figure 3.1: An image of the hypothetical learning trajectory used in this study. The top row represents evolutions in definitions, the second row represents changes in dominant van Hiele level, and the third level represents changes in conceptions of shape types.

interactions between them. The hypothetical learning trajectory created serves as a potential basis for additional studies which combine multiple theoretical perspectives and explore multiple facets in order to achieve a more comprehensive description of the learning process.

Nearly every task used in the teaching experiment portion of the study involved composing a mathematical definition in some way. Every time a student composed a definition she was prompted to self-assess it by being asked the following questions: "Is everything in your definition true about all examples of a [concept]?" "Are there enough details in your definition to tell the difference between an example of a [concept] and something that should not count as a [concept]?" and "Do you need all of the parts that you have included in your definition in order for it to still distinguish between examples and non-examples?" these question were chosen to highlight the need for necessary, sufficient, and minimal conditions, respectively.

Results

The results of the initial survey revealed that the majority of students were not capable of consistently composing definitions with necessary, sufficient and minimal conditions at the beginning of the study. A more detailed account of how students performed with definitions of each quadrilateral type can be found in Table 1.

Based on the survey results, three students were chosen to participate in the teaching experiment portion of the study. These students represented each of the following cases, a student who knew relatively little about mathematical definitions, a student who wrote definitions which were lengthy but accurate descriptions, and a student consistently wrote definitions with necessary, sufficient, and minimal characteristics. All three students showed improvement in their understanding of mathematical definitions in some way. They all began anticipating and adjusting for the self-assessment questions addressing necessary, sufficient, and minimal conditions. This generally improved their defining ability and showed that they understood and valued these qualities of a mathematical definition. By the end of the study all three students included some equivalent description of definitions as having necessary, sufficient, and minimal conditions on their final survey.

Improvements were evident in various ways for each student including the kind of definitions they wrote and the way they described definitions in general. The first student greatly improved her ability to write definitions with necessary and sufficient and sometimes minimal conditions improving from four necessary, no sufficient, and no minimal definitions,

Table 1

The Number (and Percent) of Definitions Which Contained Necessary, Sufficient, and Minimal Sets of Conditions by Quadrilateral Type (N=44)

Quadrilateral Type	Necessary	Sufficient	Minimal
Quadrilateral	27 (61%)	27 (61%)	20 (45%)
Square	44 (100%)	22 (50%)	14 (32%)
Parallelogram	39 (89%)	18 (41%)	10 (23%)
Trapezoid	13 (30%)	9 (20%)	6 (14%)
Rectangle	41 (93%)	16 (36%)	4 (9%)
Kite	10 (23%)	3 (7%)	3 (7%)
Rhombus	30 (68%)	10 (23%)	2 (5%)
Totals (N=308)	204 (66%)	105 (34%)	59 (19%)

to seven necessary, five sufficient, and four minimal on a set of definitions for seven terms. The second student acknowledged the importance of minimalness and was now able to more consistently compose definitions with this characteristic. It appeared that the third students' initial desirable set of definitions resulted more from memorized fact and limited knowledge of the concepts being defined rather than from a mastery of the concept of definitions. By the end of the study she retrogressed in her ability to write definitions with minimal conditions but increased her knowledge and understanding of definitions in general as evidenced by her ability to identify but not correct the shortcomings of her final set of definitions.

Implications

Although this particular study was conducted in the context of elementary education majors and basic geometry, the results provide implications which can be adapted for teaching definitions to many different audiences and in a variety of content areas. It supports de Villiers, Govender and Patterson's (2009) proposal that the act of composing mathematical definitions will increase students' understanding of mathematical definitions in general. This suggests that writing definitions can be a productive and worthwhile endeavor for any doers of mathematics who does not fully understand this construct. This can be especially true for undergraduates who may never have developed a complete understanding of definition on the high school level and are about to encounter sophisticated mathematical definitions and related proofs at the undergraduate level.

The results also inform undergraduate mathematics educators with regard to the role of students' previous knowledge. Theories such as van Hiele theory (Burger & Shaughnessy, 1986) and the Conceptual Learning and Development Model (Sowder, 1980) rely on a student's ability to compose a mathematical definition as evidence of a certain level of understanding of the concept being defined. However some students, like the third student in this study, may be able to provide accurate definitions as a result of previous exposure and memorization rather than as a result of conceptual understanding. This highlights the importance of always reviewing concepts and verifying students understanding rather than assuming understanding based on the use of definitions alone.

References

Burger, W., & Shaughnessy, J. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for Research in Mathematics Education, 17*(1), 31–48.

Chesler, J. (2012). Preparing mathematics teachers to teach definitions. In S. Brown, S. Larsen, K. Marrongelle, & M. Oehrtman (Eds.), *Proceedings of the 15th annual conference on research in undergraduate mathematics education* (pp. 39-44). Portland, OR: Authors.

Confrey, J., & Lachance, A. (2000). Transformative teaching experiments through conjecture driven research design. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education* (pp. 231-266). Hillsdale, NJ: Erlbaum.

de Villiers, M., Govender, R., & Patterson, N. (2009). Defining in geometry. In T. V. Craine & R. Rubenstein (Eds.), *Understanding geometry for a changing world* (pp. 189-203). Reston, VA: National Council of Teachers of Mathematics.

Fujita, T. (2012). Learners' level of understanding of the inclusion relations of quadrilaterals and prototype phenomenon. *Journal of Mathematical Behavior, 31*(1), 60-72.

Leikin, R., & Winicki-Landman, G. (2001). Defining as a vehicle for professional development of secondary school mathematics teachers. *Mathematics Education Research Journal, 3*, 62-73.

Linchevsky, L., Vinner, S., & Karsenty, R. (1992). To be or not to be minimal? Student teachers' views about definitions in geometry. *Proceedings of PME, 16*, 48-55.

Sowder, L. (1980). Concept and principle learning. In R. L. Shumway (Ed.), *Research in mathematics education* (pp. 244-285). Reston, VA: National Council of Teachers of Mathematics.

Usiskin, Z., & Griffin, J. (2008). *The classification of quadrilaterals: A study of definition.* Charlotte, NC: Information Age Publishing.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 65-81). Dordrecht, The Netherlands: Kluwer Academic Publishing.

Zaslavsky, O., & Shir, K. (2005). Students' conceptions of a mathematical definition. *Journal for Research in Mathematics Education, 36*(4), 317-346.

PRESERVICE SECONDARY TEACHERS' UNDERSTANDING OF THE CARTESIAN CONNECTION AND EQUIVALENCE

Kyunghee Moon
University of West Georgia

Both prior research and national standards emphasize the importance of critical ideas, such as the Cartesian Connection and equivalence, in algebra problem solving. The mathematics education community, however, has yet to determine whether the secondary teachers who teach such ideas fully grasp these ideas themselves. To investigate this, I interviewed a cohort of nine preservice secondary teachers in a teacher education program with two algebra problems that embed these ideas. The results showed that many of the teachers failed to understand equivalence as a relation between geometric objects, and thus could not solve an algebra problem by relating algebraic equations to their corresponding graphs. Many also misinterpreted the meaning of the term "solution," and thus could not use the Cartesian Connection to find a solution of an equation. It is advisable that secondary teacher education programs focus more on these critical ideas so that secondary teachers can impart such ideas on their students.

Keyword: Cartesian Connection, Equivalence, Representations, Preservice secondary teachers

Introduction

This study investigates preservice secondary teachers' content knowledge, and in particular their ability to connect algebraic and graphical representations in algebra problem solving. The Principles and Standards for School Mathematics (NCTM, 2000) and the Common Core State Standards (CCSS) emphasize the role of representations in mathematical teaching and learning. Prior research also suggests that connecting representations is vital for successful problem solving or mathematical understanding (Brenner et al., 1997; Hiebert & Carpenter, 1992; Tompson, 1994). Many of these studies have showed that connecting representations is not trivial even for those who have advanced mathematical training (Even, 1998; Gagatsis & Shiakalli, 2004; Knuth, 2000).

Studies (Kieran, 1981; Knuth, 2000; Moschkovich, Schoenfeld, & Arcavi, 1993; You, 2005) have suggested that the failure to understand critical ideas in mathematics, such as the Cartesian Connection: "a point is on the graph of the line L if and only if its coordinates satisfy the equation of L" (Moschkovich, Schoenfeld, & Arcavi, 1993, p. 73), is a main cause of students' difficulties in connecting representations. Studies have also documented that many teachers do not understand students' difficulties with these critical ideas (Postelnicu, 2011). Due to lack of research, however, the mathematics education community does not understand whether secondary teachers themselves (who are, or will be, teaching these ideas) are indeed equipped with full understanding of these critical ideas. In this research, I examine preservice secondary teachers' knowledge of these ideas through the use of interviews. In particular, I focus on two of the most discussed ideas in algebra—the Cartesian Connection and equivalence.

Conceptual Framework

This study examines preservice teachers' understanding of "big ideas"—the Cartesian

Connection and equivalence. Big ideas are "central organizing ideas of mathematics—principles that define mathematical order" (Schifter & Fosnot, 1993, p. 35). These ideas are considered "big" because "they are critical to mathematics and because they are big leaps in the development of children's reasoning" (Fonsnot & Jacob, 2010, p. 29). The relationship between one's problem solving ability and translation ability among representations is documented in research (Brenner et al., 1997; Hitt, 1998; Thompson, 1994). The association of big ideas to one's translation ability is also documented in prior research, as shown in Moschkovich et al. (1997) and Knuth (2000), with the big idea of the Cartesian Connection. As such, individuals' problem solving abilities are closely related to their understanding of big ideas.

The two big ideas this study focuses on are the Cartesian Connection and equivalence in the topic of algebra. The Cartesian connection is an idea that connects algebraic and graphical representations relating a point on the graph of an equation and its coordinates as a solution of the equation. Equivalence is an idea that has been previously discussed mostly within the algebraic context, but is extended in this study as an idea that relates algebraic and graphical representations.

Research has showed that many secondary students and K-8 teachers do not firmly grasp the idea of the Cartesian Connection, and thus use algebraic approaches in problem solving when geometric approaches are more efficient. For instance, Knuth (2007) showed that more than three-fourths of the study's participants (178 high school students) used algebraic approaches to find a solution of the algebraic equation, $?x + 3y = -6$, when a solution could be found simply by inspecting the given graph. You's study (2006) also showed that only 27 out of 104 elementary and middle school teachers were able to provide a solution correctly for the same question Knuth used. Among those teachers who provided a correct solution, only 7 teachers used a geometric approach. The study by Moon et al. (2013) also showed prospective secondary teachers' difficulties in applying the Cartesian Connection in an inquiry-based, problem solving class. Moon et al. extended the definition of the Cartesian Connection to "A point is on the graph of the mathematical relation $R(x,y)=0$ if and only if its coordinates satisfy $R(x,y)=0$" (p. 204) to discuss the idea in the context of conic curves.

Equivalence is the idea that 'the equal sign denotes an equivalence relation between two quantities' (Kieran, 1981; Knuth, Stephens, McNeal, & Alibali, 2006). Most students in primary grades hold an operational view of the equal sign, and thus provide an answer such as 12 or 17 for the question asking the missing value in $4 + 3 + 5 = \_\_ + 5$, by performing operations with the numbers involved in the equation (Alibali, 1999). Many middle grade students also possess this operational view and have difficulty finding the value of m using appropriate algebraic approaches when solving equations such as $4m + 10 = 70$ (Knuth et al., 2006). Many high school or college students also do not use the equal sign correctly in problem solving, providing a mathematical statement such as $x + 3 = 7 = 7 – 3 = 4$ (Kieran, 1981).

Although this idea of equivalence is treated in research and practice mostly as a relation between two algebraic objects, it is a big idea that relates algebraic and graphical representations and that leads to successful problem solving in algebra. For example, in order for a student to explain an Algebra standard: "why the x-coordinates of the points where the graphs of the equations $y = f(x)$ and $y = g(x)$ intersect are the solutions of the equation $f(x) = g(x)$" (CCSS, p. 66), the student needs to understand equivalence as a relation between two geometric objects, along with other big ideas such as the Cartesian Connection and 'the graph of an equation as the set of points whose coordinates satisfy the equation'. To be more specific, the student should be able to see the equal sign in the algebraic equation, $f(x) = g(x)$, as a relation between two algebraic objects, $f(x)$ and $g(x)$, and at the same time should be able to connect the algebraic objects to geometric objects—the graphs of functions, $y = f(x)$ and $y = g(x)$. The student also needs to understand that the intersection points of the two graphs, which could be denoted by (a, b) (representing unknown points), indeed lie on both graphs, and as such the pair, $x = a$ and $y = b$, has to satisfy both equations, $y = f(x)$ and $y = g(x)$, applying the idea of the Cartesian Connection. The student then needs to understand that the relations $b = f(a) = g(a)$ explain $x = a$ as the solution of the equation $f(x) = g(x)$. Yet, as illustrated in Dufour-Janvier (1987) with a case of one high school student, an individual who only sees $f(x) = g(x)$ as an algebraic relation might not be able to accept the fact that the solutions of the equation $f(x) = g(x)$ can be explained by the intersection points of the graphs of $y = f(x)$ and $y = g(x)$, even after observing the graphs of the equations.

As seen above, both the Cartesian Connection and equivalence as a relation between geometric objects are big ideas that are crucial to successful problem solving, the ideas that teachers need for secondary teaching. Thus, in this research, I study a cohort of preservice secondary teachers to answer the research question: "What is preservice secondary teachers' understanding of the Cartesian Connection and equivalence as shown in algebra problem solving?"

Methods

This study was conducted at a large public university located in southern California. In the 2010-2011 school year, 14 preservice secondary mathematics teachers (PSMTs) entered the teacher education program to pursue secondary teaching credentials in mathematics. The teacher education program was a one-year graduate program, and as such, all PSMTs held a bachelor's degree in mathematics, science, or engineering. The PSMTs took a mathematics methods course in the fall quarter, mathematics pedagogy courses in the winter and spring quarters, and a mathematics problem solving course in the spring quarter. Both methods and problem solving courses emphasized the importance of connecting representations in mathematical problem solving and understanding. However, there was no specific instruction on the Cartesian Connection or equivalence in these courses.

The data source used in this study is a portion of data used for larger research that dealt with PSMTs' mathematical knowledge for teaching. Specifically, it was from a one-hour interview, designed to examine PSMTs' abilities to connect representations, which was conducted during the winter quarter. Out of 5 interview questions, 2 questions concerned the Cartesian Connection and/or equivalence (see Figure 1). Only 9 of the 14 PSMTs participated in the interview. During the interview, PSMTs were asked to "think aloud" so that their thought process could be visible through their words. They were also asked to do mathematical work on the interview sheets. The interview was videotaped using a video camera and was transcribed before analysis.

Figure 1 shows the two questions used in this study. The first question Q1 was designed, inspired from Dufour-Janvier et al. (1987), so that the three subquestions (A), (B) and (C) would measure PSMTs' understanding of equivalence as a relation between algebraic objects, equivalence as a relation between geometric objects, and the Cartesian Connection, respectively. The second question Q2 was quoted straight out of Knuth (2000) to examine PSMTs' understanding of the Cartesian Connection.

Q1. Cynthia has a hard time in figuring out the solution(s) to the equation $\sqrt{x} = x - 2$ (Dufour-Janvier et al., 1987).

(A) How would you help her?
(B) How could you use graph(s) to help her understand?
(C) How is the solution of the equation $\sqrt{x} = x - 2$ related to the graphs in (B)?

Q2. The graph below represents the equation, $?x + 3y = -6$. (We do not know the value of the coefficient of x.) (Knuth, 2000)

(A) Is it possible to find a solution to the equation without the missing value? Explain your answer.

(B) How could we find the missing value? Explain your answer.

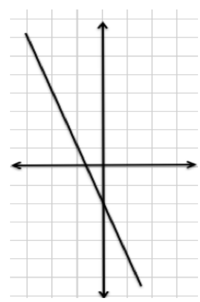


Figure 1. Interview Questions

For analysis, I coded PSMTs' written work and the transcription of their interviews. For Q1, I coded PSMTs' problem solving abilities in each of the three subquestions: to find the solution of $\sqrt{x} = x - 2$ algebraically from Q1(A), to connect the equation $\sqrt{x} = x - 2$ to its corresponding graphs from Q1(B), and to connect the intersection points of graphs back to the solution of the equation from Q1(C).

For Q2, I coded PSMTs' problem solving strategies, similar to Knuth (2000). For Q1(A), I examined whether the participants used the given graph (without solving the equation) or the algebra equation as the main means to find a solution. For Q2(B), I coded whether PSMTs substituted the x and y coordinates of a point from the graph directly into the equation, $?x + 3y = -6$, to look for the missing value, or whether they found the slope of the equation first and then used the slope to look for the missing value.

Results

Results of the analysis are presented here, first for Question 1 and then for Question 2. For Question 1, I explain PSMTs' understandings of equivalence as an algebraic relation and of equivalence as a geometric relation, along with their understanding of the Cartesian Connection.

For Question 2, I explain their understanding of the Cartesian Connection and an unexpected issue that PSMTs encountered in problem solving.

Q1: The Solution of $\sqrt{x} = x - 2$

Equivalence as an algebraic relation. When asked to find the solution of the equation, $\sqrt{x} = x - 2$ (Q1(A)), all nine PSMTs used algebraic approaches. Eight PSMTs squared both sides of the equation to convert the equation to $x^2 - 5x + 4 = 0$ and found the solutions, 1 and 4, of the resulting equation, $x^2 - 5x + 4 = 0$. Four of those eight PSMTs followed up by inputting the numbers in the original equation $\sqrt{x} = x - 2$ and then claimed that 4 was the only solution. The other four PSMTs claimed that both 1 and 4 were the solutions with no further work. When asked if both were the solutions, however, two of those four PSMTs input the numbers in the original equation and revised their answers to state that 4 was the only solution. The other two PSMTs checked the numbers by using the equation $x^2 - 5x + 4 = 0$ instead, and claimed that both 1 and 4 were the solutions of the equation $\sqrt{x} = x - 2$. When asked if checking with $x^2 - 5x + 4 = 0$ was a valid method, those two PSMTs answered that it was valid because $x^2 - 5x + 4 = 0$ was derived from $\sqrt{x} = x - 2$. These two had a problem with understanding logical equivalence: 'if $a = b$ then $a^2 = b^2$, but $a^2 = b^2$ is not an equivalent statement to $a = b$'. Two PSMTs, including one of the four PSMTs who solved it correctly above, used a different algebraic approach. They first changed the equation $\sqrt{x} = x - 2$ to $x - \sqrt{x} - 2 = 0$ and then factored the second equation into $(\sqrt{x} - 2)(\sqrt{x} + 1)$. They then claimed that 4 was the only solution of the equation as there is no real value of x that makes $\sqrt{x} = -1$. It seemed that all nine PSMTs had no problem seeing the equal sign in $\sqrt{x} = x - 2$ as an algebraic relation, but some of them had difficulty with this question due to their lack of understanding in logical equivalence.

Equivalence as a geometric relation and the Cartesian Connection. In order to investigate PSMTs' understanding of equivalence as a geometric relation and of the Cartesian Connection in Q1, I categorized the participants according to their problem solving abilities in Q1(B) and Q1 (C). The vertical side of Table 1 is for Q1(B). If a participant related graphs to the equation $\sqrt{x} = x - 2$ on their own, I categorized the participant into "Using graph with no assistance." If not, I categorized the participant into "Using graph with assistance." The horizontal side of Table 1 is for Q1(C). If a participant interpreted the x-coordinate(s) of the intersection point(s) of the graph(s) as the solution(s) of the equation, I categorized the participant into "Able to connect the solution to the intersection points." If not, I categorized the participant into "Unable to connect the solution to the intersection points."

Table 1
Problem Solving in Q1(B) and Q1(C)

	Q1(C)	
Q1(B)	Able to connect the solution to the intersection points	Unable to connect the solution to the intersection points
Using graph with no assistance	4 participants (Cell 1)	1 participant (Cell 2)
Using graph with assistance	4 participants (Cell 3)	0 participant (Cell 4)

Interpreting the results of the categorizations, four PSMTs (Table 1, Cell 1) showed an understanding of equivalence as a geometric relation and of the Cartesian Connection. Three of these four PSMTs related $\sqrt{x} = x - 2$ to functions, $y = \sqrt{x}$ and $y = x - 2$, and sketched the graphs of the functions (equivalence as a geometric relation). They then claimed that the equation $\sqrt{x} = x - 2$ had only one solution 4 as the two graphs met at one point (the Cartesian Connection). These three PSMTs were in fact among those four PSMTs who solved Q1(A) correctly by using an algebraic approach. One PSMT sketched a graph of $y = x^2 - 5x + 4$ instead of graphs of $y = \sqrt{x}$ and $y = x - 2$. This PSMT was one of the two who had come up with two solutions for Q1(A), believing that $\sqrt{x} = x - 2$ embedded the same information as $x^2 - 5x + 4 = 0$. Although she worked with the incorrect equation, her work showed that she had an understanding of both equivalence as a geometric relation and the Cartesian Connection.

One PSMT (Table 1, Cell 2) showed an understanding of equivalence as a geometric relation, but not of the Cartesian Connection. This PSMT sketched graphs of $y = \sqrt{x}$ and $y = x - 2$ on her own. However, with the graph of $y = \sqrt{x}$ vastly skewed, her two graphs did not intersect. She was not bothered by this conflicting information—no intersecting points even though she had already found one solution of the equation using an algebraic approach. With my assistance, she fixed her graphs with the two graphs meeting at one point. However, when asked to interpret the solution of the algebraic equation $\sqrt{x} = x - 2$ related to the graphs, she said, "if there are two statements, there is a solution. But I am not sure if that is the intersection." I asked if the solution $x = 4$ was related to the intersection point, and she answered "not necessarily." It was obvious that her knowing that the number of solutions of the equation $f(x) = g(x)$ is the number of intersection points in the graphs $y = f(x)$ and $y = g(x)$ was from her memorization of the rule, and not from an actual understanding of the Cartesian Connection.

Four PSMTs (Table 1, Cell 3) did not show an understanding of equivalence as a geometric relation, but did show an understanding of the Cartesian Connection. When they were asked to explain the solution of $\sqrt{x} = x - 2$ using graphs, their replies were, "how can I make this as *y* equals something like *x* to have input and output values? There is no *y*," "every time when somebody says to draw a graph, it is in *y* equals form. But when I see this, there is no *y* or $f(x)$ or anything like that," "I just can't add *y*," and "I know how to graph $y = \sqrt{x}$, but not $\sqrt{x} = x - 2$." However, when prompted to sketch the graphs of the equation $y = \sqrt{x}$ and $y = x - 2$, all four PSMTs were able to sketch both graphs correctly (with one PSMT needing some assistance). When they were asked to interpret the solution of the equation using graphs, all of them claimed that $x = 4$ was the unique solution of the equation $\sqrt{x} = x - 2$ by referring to the intersection point of the graphs (the Cartesian Connection).

Q2: The Solution of $?x + 3y = -6$

The Cartesian Connection. Before interpreting PSMTs' understanding of the Cartesian Connection in the problem solving of Q2, I will first explain how Knuth (2000) interpreted high school students' understanding of the Cartesian Connection where the students were examined by a paper and pencil test. Knuth categorized students' problem solving approaches as either geometric or algebraic. If a student used the given graph as the main means to find the solution—for instance, visually inspecting the graph to find the answer—, Knuth categorized the approach as geometric. If, however, a student used an algebraic procedure as the main means, such as finding the slope of the line and using the slope algebraically, he categorized the approach as algebraic. Knuth reported that more than three-fourths of the high school students in his study used algebraic approaches and that more than 90% of the students who used algebraic approaches found the slope of the line as part of their solution methods. He claimed that "students' reliance on algebraic-solution methods was due to their failure to recognize the points used in calculating a slope as solutions to an equation," which explained the lack of understanding of the Cartesian Connection (p. 505).

Using interviews as my research instrument (although using the same questions as Knuth did), I encountered an unexpected situation: While solving question Q2(A), six of the PSMTs asked or confirmed with me the meaning of the term "solution" in the context of the problem. When I asked them what they believed the "solution" meant, all of them replied that the "solution" would mean the missing value, ?. In order to help them move forward, I told them that "solution" in the context of the problem meant x and y values, information which affected their problem solving approaches, and hence my analysis of their responses.

Table 2 shows my categorization of PSMTs' responses to Q2(A) and Q2(B). The vertical side of Table 2 shows their responses to Q2(A). If a participant found a solution by simply inspecting the given graph and did not ask me to clarify the term "solution," the participant was categorized as "Geometric approach with no assistance." If a participant found a solution by inspecting the given graph and asked for my assistance on the meaning of "solution," the participant was categorized as "Geometric approach with assistance." If a participant used an algebraic approach for Q2(A), such as changing the equation $?x + 3y = -6$ into the slope-intercept form $y = (-?/3)x - 2$ and determining ? using the slope of the equation, the participant was categorized as "Algebraic approach" regardless of my assistance on the meaning of "solution." The horizontal side of Table 2 is the categorization of their responses to Q2(B). If a participant simply substituted the x and y coordinates of a point on the graph into the equation to find the value of the missing value, the participant was categorized as "Direct substitution." If a participant found the value of the slope and used the slope to determine the value of ?, the participant was categorized as "Using slope."

Table 2
Problem Solving Approaches in Q2(A) and Q2(B)

	Q2(B)	
Q2(A)	Direct substitution	Using slope
Geometric approach with no assistance	2 participants (Cell 1)	0 participant (Cell 2)
Geometric approach with assistance	3 participants (Cell 3)	2 participants (Cell 4)
Algebraic approach	0 participant (Cell 5)	2 participants (Cell 6)

Interpreting the results from the categorizations in Table 2, only two PSMTs in Cell 1 (Table 2) showed a sound understanding of the Cartesian Connection, as they used efficient problem approaches based on the Cartesian Connection in both Q2(A) and Q2(B). These two PSMTs would surely have solved both problems using geometric approaches even in a paper-and-pencil environment as in Knuth (2000). For the next group of three PSMTs in Cell 3 (Table 2), I interpret that these PSMTs also had a sound understanding of the Cartesian Connection, as they were able to use efficient approaches in both Q2(A) and Q2(B). However, it is not predictable how they would have responded to those questions had they been tested in a paper-and-pencil format, as they were confused about the meaning of the term "solution." The two PSMTs in Cell 4 had partial understanding of the Cartesian Connection. Although they used a geometric approach in Q1(A), they used an inefficient approach in Q2(B)—finding the slope of the line to look for the missing value—which could be solved much more simply had they used the Cartesian connection. This group surely would have been placed in the group showing lack of understanding of the Cartesian connection in Knuth's study. The last group, two PSMTs in Cell 6, most certainly showed a lack of understanding of the Cartesian Connection, both in my analysis and Knuth's. One PSMT used inefficient approaches after consulting with me about the meaning of "solution" and the other did so without consultation.

Discussion

The results of this study indicate that many preservice secondary teachers have difficulty applying in problem solving the ideas of the Cartesian Connection and equivalence as a geometric relation. They also suggest that many have difficulties interpreting the term "solution" when an equation has both variables and an unknown value. On the question that asked for a "solution" of the equation $?x + 3y = -6$ and the value of the question mark, only two out of nine PSMTs interpreted the word "solution" correctly and showed a sound understanding of the Cartesian Connection without any assistance. Three of them showed a sound understanding of the Cartesian Connection with my assistance on the meaning of the term "solution," while four of them showed a lack of understanding of the Cartesian Connection. Considering the fact that about a quarter of high school students in Knuth (2000) used geometric approaches for the same question with no assistance (as they were tested in a paper and pencil format), the preservice teachers in this study did not perform any better than the high school students in Knuth.

In the question that asked for a geometric interpretation of the solution of the equation $\sqrt{x} = x - 2$, only four out of nine PSMTs were able to interpret successfully, relating the equation to functions and the solution of the equation to the graphs. Out of those five preservice teachers who could not successfully interpret the solution geometrically, four of them could not relate any graphs to the equation $\sqrt{x} = x - 2$ (showing a lack of understanding of equivalence as a relation between geometric objects). The remaining one PSMT was able to relate the equation to the graphs of the functions $y = \sqrt{x}$ and $y = x - 2$, but was unable to explain that the solution of the equation was the x-intercept of the intersection points of two graphs (showing a lack of understanding of the Cartesian Connection). Dufour-Janvier et al. (1987) discussed this problem with a case of a high school student who could only make sense of the solution of the equation through algebraic procedures. Unlike the student in Dufour-Janvier, none of the preservice teachers in this study had a problem accepting the fact that the solution of the equation was related to the graphs of the corresponding functions, once they were told that the graphs of the functions, $y = \sqrt{x}$ and $y = x - 2$, are related to the equation $\sqrt{x} = x - 2$.

This study suggests a few implications for secondary education and secondary teacher education. First, the mathematics education community needs to understand the complexity involved in many national standards such as CCSS. For instance, the CCSS simply states that students need to be able to "explain why the x-coordinates of the points where the graphs of the equations $y = f(x)$ and $y = g(x)$ intersect are the solutions of the equation $f(x) = g(x)$" (p. 66), without specifying what ideas or skills are involved in this process. As shown in this study with the example of $\sqrt{x} = x - 2$, successfully performing such a task is not trivial. A student needs to understand the meaning of the problem's "solution," solve it algebraically, see $\sqrt{x} = x - 2$ as a relation between geometric objects—the graphs of the functions, $y = \sqrt{x}$ and $y = x - 2$, correctly transfer the function equations $y = \sqrt{x}$ and $y = x - 2$ to their graphs, and interpret the intersection point of the graphs connected to the equation $\sqrt{x} = x - 2$. Without unpacking the core ideas and skills behind the standards, it would be very difficult to help students meet the standards.

This study also suggests that secondary teachers might not be fully equipped with the critical mathematical ideas that they need for teaching, an implication that dampens the community's hopes that secondary teachers with advanced mathematics courses have sound understanding of content, at least in algebra. These critical ideas were cognitive blocks for successful problem solving not only for students (Dufour-Janvier, 1987; Kieran, 1981; Knuth, 2000), but also for secondary teachers, as seen in this study. Teachers' knowledge of content influences their teaching and/or student learning (Borko & Eisenhart, 1992; Stein, Baxter, & Leinhardt, 1990). As such, it is important that these secondary teachers are provided with opportunities to develop such knowledge through mathematics courses and professional development programs.

There is an urgency to improve students' understanding of mathematics at the secondary level. Without the development of mathematical concepts/ideas and reasoning skills, there is only a slim chance for students to succeed both academically and professionally in this rapidly changing world. There is still so much unknown about how to better prepare secondary teachers and how to better help students learn mathematics. It is time for the field to pay more attention to critical issues in secondary education.

References

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durán, R., Smith Reed, B., & Webb, D. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal*, *34*(4), 663-689.

Borko, H., & Eisenhart, M. (1992). Learning to teach hard mathematics: Do novice teachers and their instructors give up too early? *Journal for Research in Mathematics Education*, *23* (3), 194-222.

Broody, J., Feil, Y., & Jhonson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for research in Mathematics Education*, *38*(2), 115-131.

CCSS (Common Core State Standards). http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Even R. (1998). Factors involved in linking representations of functions. *Journal of Mathematical Behavior, 17*(1), 105-121.

Carpenter, T. P., Fennema, E., & Peterson, P. L. (1988). Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education*, *19*(5), 385-401.

Gagatsis, A. & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology*, *24*, 645-657.

Hiebert, J ., & Carpenter, T. P. (1992). Learning and teaching with understanding. n D. A. Grouws(Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.
Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38*(3), 258-288.
Kieran, C. (1981). Concept associated with equality symbol. *Educational Studies in Mathematics*, *12*(3), 317-326.

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 707–62). Charlotte, N.C.: Information Age Publishing; Reston, Va. National Council of Teachers of Mathematics.
Knuth, E. (2000). Student understanding of the Cartesian coordinate connection: An exploratory study. *Journal for Research in Mathematics Education*, *31*(4), 500-508.
Knuth, E. J., Stephen, A.C., McNeil, N.M., & Alibali, M.W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*, 297-312.
Moschkovich, J., Schoenfeld, A. H., & Arcavi, A. (1993). Aspects of understanding: On multiple perspectives and representations of linear relations, and connections among them. In T. Romberg, E. Fennema, & T. Carpenter (Eds.), *Integrating research on the graphical representation of function* (pp. 69–100). Hillsdale, NJ: Erlbaum.
National Council of Teachers of Mathematics (NCTM, 2000). *Principles and standards for school mathematics*. Reston, VA: Author.
Stein, M. K., Baxter, J. A., & Leinhardt, G. (1990). Subjectmatter knowledge and elementary instruction: A case from functions and graphing. *American Educational Research Journal, 27*, 639-663.
Thompson, Patrick W. (1994). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics*, *26*, 229-274.
Yerusalmy, M., & Schwartz, J. L. (1993). Seizing the opportunity to make algebra mathematically and pedagogically interesting. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating Research on the Graphical Representation of Functions* (pp. 41-68).

103

WHAT CONSTITUTES A WELL-WRITTEN PROOF?

Robert C. Moore
Andrews University

The purpose of this study was to identify some of the characteristics mathematicians value in good proof writing. Four mathematicians were interviewed. First, they evaluated and scored six proofs of elementary theorems written by students in a discrete mathematics or geometry course, and second, they responded to questions about the characteristics they value in a well-written proof and how they communicate these characteristics to students. Preliminary results indicate that these mathematicians agreed that the most important characteristics of a well-written proof are (a) correct logic and (b) clarity. Although these mathematicians differed in the attention they give to layout, grammar, punctuation, and mathematical notation, they agreed in giving these characteristics relatively little weight in the overall score. The results also showed that, in addition to demonstrating good proof writing in class, writing comments on students' papers is an important way they teach their students to write good proofs.

Key words: Proof writing, Proof evaluation, Proof assessment, Teaching proof

Introduction and Related Research

This study is an initial exploration into mathematicians' beliefs and practices in evaluating students' written proofs. Speer, Smith, Horvath (2010) have called for research on how teachers of undergraduate mathematics design assessments, what they value in students' responses, and how they communicate those values and evaluate students' work. Inglis, Mejía-Ramos, Weber, and Alcock (2012) also noted that "mathematicians' grading and instruction on what constitutes a proof is a useful avenue for future research" (p. 70). The present study is a first step toward addressing these calls for research on university teachers' proof assessment practices.

As identified by Mejía-Ramos and Inglis (2009), one proof-related activity in which undergraduate students engage is that of demonstrating understanding by constructing and presenting an argument to an expert, namely, their teacher. A complementary activity, performed by the teacher, is that of evaluating a student's argument against a set of criteria. While considerable attention has focused on students' difficulties and errors in proof construction (e.g., Harel & Sowder, 1998; Moore, 1994; Selden & Selden, 1995; Weber, 2001), little attention has been given to how mathematicians respond to students' errors.

One aspect required in the evaluation of a student's argument is that of deciding whether the argument is indeed a valid proof. Weber (2008) found disagreement among mathematicians about the validity of purported proofs, as did Inglis and Alcock (2012) who concluded that "some of these disagreements could be genuinely mathematical, rather than being relatively trivial issues related to style or presentation" (p. 379). In their study of 109 research-active mathematicians, Inglis, Mejía-Ramos, Weber, and Alcock (2013) found further evidence that there is no universal agreement about what constitutes a valid proof. They concluded that mathematicians' standards of validity differ and questioned whether students are getting a consistent message about what constitutes a valid proof.

In order to better understand the features that mathematicians value in proofs written for pedagogical purposes, Lai, Weber, Mejía-Ramos (2012) observed mathematicians as they revised two proofs. They found that mathematicians often agree, but occasionally disagree to a remarkable extent, on whether a proposed revision of a proof will improve the proof or make it worse for pedagogical purposes. Note that this study focused on proofs written by mathematicians for presentation to students. I am not aware of any empirical studies that have examined the characteristics that mathematicians value in students' proof writing.

Given that mathematicians sometimes differ on the question of validity and other features of a written argument, it is reasonable to expect that they may differ in their assessment of students' work, which requires judgment calls not only on validity, but also on clarity, readability, and other features. Furthermore, in scoring proofs professors must make decisions about partial credit. Thus, this evaluation process, which appears to be more complex than simply deciding whether an argument is valid, deserves attention.

After carefully examining students' papers, Brown and Michel (2010) created an assessment rubric built on three characteristics: readability, validity, and fluency (RVF). The authors claimed that it aids in the efficiency and consistency of evaluating students' work, serves as a means of communicating to students the characteristics of good writing, and provides feedback for improvement. The question arises as to whether other mathematicians would agree with this list of characteristics of good mathematical writing. Given a student's proof, would they evaluate it the same way?

Research Questions

These studies motivated the following main questions of this study:

1. Do mathematics professors agree in their evaluation and scoring of students' proofs?
2. What characteristics do mathematics professors consider when evaluating students' proofs?

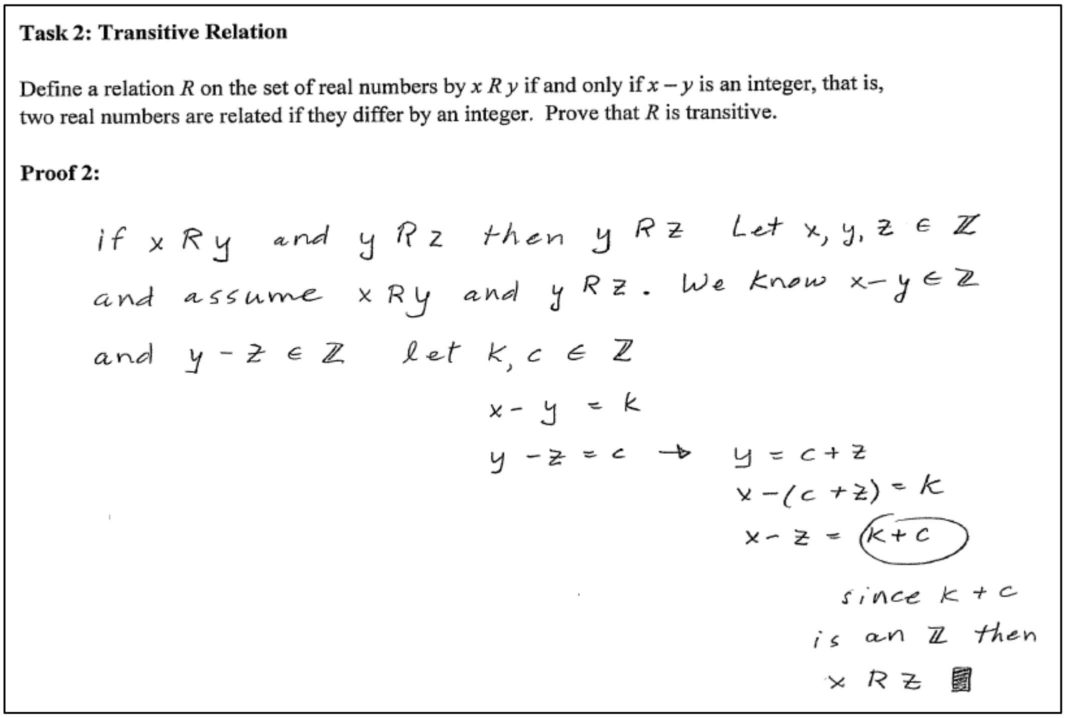
Methodology

I conducted individual interviews that lasted about one hour with four mathematics professors, three women and one man, all from a small private university. Each professor had a Ph.D. in mathematics, was actively involved in research, and had at least a dozen years of university-level mathematics teaching experience, including advanced undergraduate courses. Three of them had taught proof-based courses such as linear and abstract algebra, advanced calculus, and complex variables. The fourth professor primarily taught applied mathematics but emphasized proofs in her calculus courses and had taught an introduction to proof course once.

In the first part of the interview, the participants talked aloud as they evaluated five or six written proofs by (a) marking on the proof what was correct or incorrect, (b) telling how the proof could have been improved, and (c) assigning a score out of ten points to the proof. In the second part of the interview, the participants responded to questions about their beliefs and practices in evaluating students' proofs and teaching students to write proofs.

The six proofs used for the first part of the interview were proofs of elementary theorems written by my students for tests or homework: five proofs from a discrete mathematics course, which serves as a transition-to-proof course for mathematics majors, and one proof from a geometry course. I chose them because they contained a variety of features that I judged to be both good and bad that related to readability, validity, fluency, proof frameworks, clarity of reasoning, use of definitions, quantifiers, and surface features such as layout and punctuation (Brown & Michel, 2010; Lai et al., 2012; Moore, 1994; Selden & Selden, 1995). As an example, Figure 1 shows one of the six proofs.

The interviews were recorded with a Livescribe smart pen that recorded the participants' handwriting on the proofs as well as their talk-aloud commentary. I transcribed the interviews and analyzed them using an open coding system (Strauss & Corbin, 1990). I began the analysis



Task 2: Transitive Relation

Define a relation R on the set of real numbers by $x\ R\ y$ if and only if $x - y$ is an integer, that is, two real numbers are related if they differ by an integer. Prove that R is transitive.

Proof 2:

if $x R y$ and $y R z$ then $y R z$ Let $x, y, z \in \mathbb{Z}$

and assume $x R y$ and $y R z$. We know $x - y \in \mathbb{Z}$

and $y - z \in \mathbb{Z}$ let $k, c \in \mathbb{Z}$

$x - y = k$

$y - z = c \rightarrow y = c + z$

$x - (c + z) = k$

$x - z = k + c$

since $k + c$ is an $\mathbb{Z}$ then $x R z$ ∎

Figure 1. Task 2 and Proof 2

by making a detailed chart for each proof that identified all the written marks, corrections, and scores. Next, from the written marks, oral commentary, and responses to the interview questions, I developed codes that revealed the participants' justifications for the marks and scores and the characteristics to which they attend when evaluating proofs.

Results

Research question 1. Table 1 shows considerable variation in the scores assigned to the proofs by the four professors. (Due to time constraints, only three professors evaluated Proof 6.) While space limitations do not permit a detailed discussion of the professors' evaluations of the six proofs, Proof 2 (Figure 1) will serve as an example. All four professors focused on the proof's logic and agreed it was essentially correct, but they corrected the following errors: (a) Two noticed *y*Rz should be *x*Rz in line 1, (b) three noticed that **Z** should be **R** at the end of line 1, and (c) two said the word "let" should be "for some" in line 3. The professors also focused on clarity of reasoning and readability: (d) Three said the proof should begin with "We want to prove" and (e) three said the proof would be clearer by writing x – z = (y + k) – (y – c) = k + c in place of the work on lines 4-7. They paid little attention to punctuation and made no comment about how the proof runs diagonally down the right-hand side of the page.

The grading of this proof illustrates both agreement and disagreement in the professors' grading. They agreed that the logic of the proof was essentially correct, that the student understood the proof and its key ideas well enough, and that surface features such as layout

	Proof 1	Proof 2	Proof 3	Proof 4	Proof 5	Proof 6
Low score	8	5	6	7	7	5
High score	10	9.8	9.5	10	9.5	9.5
Mean score	9.25	7.95	8.13	8.00	7.88	7.50

Table 1. Scores assigned to the proofs by the professors

should carry little, if any, weight in the overall score. But they disagreed about how serious certain errors were, particularly the error of **Z**, rather than **R**, at the end of the first line.

A contextual factor that explains part of the spread in the scores on the six proofs is the professors' judgment about how well the student understands the proof. These professors said they base these judgments not only on the proof itself but sometimes on how well they know the student and his or her work in other situations. This finding is consistent with studies that have noted the importance of context in mathematicians' decisions on the validity of an argument (Inglis & Alcock, 2012; Inglis et al., 2013; Weber, 2008).

Research question 2. The two characteristics of a well-written proof that emerged strongly from the data were logic and clarity. The professors' comments about correct logic referred to both the overall logical structure, as well as the line-by-line reasoning. Their comments about clarity seemed to encompass a variety of meanings, including readability, justifications for the steps in the proof, and the correct use of mathematical language and notation. Here are responses from Professors C and D about the characteristics of a well-written proof:

D: Yes, the logic and clarity are the two principles. It seems like everything falls into those categories.

C: A well-written proof? The most important thing is that it's logically correct. If a proof isn't logically correct, I'll often take almost all the points off …. I give them a pretty low grade if the logic is incorrect. So that's the main thing. I would say the second thing is the readability of the proof. Is it flowing in complete sentences, or is it like vertical scratch work sort of thing?

Implications for Teaching and Further Research

I also asked the professors how they communicate the characteristics of a well-written proof to their students. Study limitations and space constraints allow me to say only that these professors use written comments on students' papers as an important means of communicating what they value in their students' proof writing. While classroom observations would help to answer this question, we should be aware that much teaching takes place outside of class.

This study shows that the assessment of students' proofs is an important means of teaching students how to write good proofs, so it behooves us, as professors, to think carefully about how we assess proofs. A rubric, such as Brown and Michel's (2010) RVF method, may serve as a means of communicating a clear, consistent message to students, thereby raising the quality of their writing throughout an undergraduate mathematics program. It is not clear to me, however, that mathematicians can agree on a general proof-writing rubric, nor whether a rubric improves consistency and efficiency in evaluating proofs, as Brown and Michel claimed. A follow-up study will delve more deeply into these questions.

Questions for the Audience

1. What suggestions do you have for a larger study of mathematicians' proof grading?
2. Is there evidence that rubrics improve professors' efficiency and consistency in evaluating students' written mathematical work?
3. Is there evidence that the use of rubrics improves students' proof writing? How could we measure this effect?

References

Brown, D. E. & Michel, S. (2010, February). *Assessing proofs with rubrics: The RVF method.* Paper presented at the 13th Annual Conference on Research in Undergraduate Mathematics Education, Raleigh, NC. Retrieved from http://sigmaa.maa.org/rume/crume2010/Archive/Brown_D.pdf

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Issues in mathematics education: Vol. 7. Research in collegiate mathematics education III* (pp. 234-282). Providence, RI: American Mathematical Society.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*, 358-390.

Inglis, M., Mejía-Ramos, J. P., Weber, K., & Alcock, L. (2012). On mathematicians' different standards when evaluating elementary proofs. In S. Brown, S. Larsen, K. Marrongelle, & M. Oehrtman (Eds.), *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 67-72), Portland, OR. Retrieved from http://sigmaa.maa.org/rume/crume2012/RUME_Home/RUME_Conference_Papers_files/RUME_XV_Conference_Papers.pdf

Inglis, M., Mejía-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science, 5,* 270-282. doi: 10.1111/tops.12019

Lai, Y. & Weber, K., Mejía -Ramos, J. P. (2012). Mathematicians' perspectives on features of a good pedagogical proof. *Cognition & Instruction, 30*(2), 146-169.

Mejía-Ramos, J. P., & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of the ICMI Study 19 conference: Proof and proving in mathematics education* (Vol. 2, pp. 88-93). Taipei, Taiwan: National Taiwan Normal University.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*, 249-266.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*, 123-151.

Speer, N. M., Smith III, J. P., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior, 29*(2), 99-114.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* London: Sage.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics, 48*, 101-119.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*, 431-459.

IMPLIED AND EMPIRICAL READERS OF NEWTON'S METHOD

Kristen Murphy
West Virginia University

Nicole Engelke Infante
West Virginia University

Celeste Glenn
West Virginia University

The ability to translate a text into a mathematical process is a key goal of mathematics education. Knowing when students have the prerequisite knowledge to understand such a process is a perennial concern for instructors. Here we use Newton's method to evaluate reader oriented theory as a means to illuminate these issues. Through clinical interviews with twelve first semester calculus students, we determined that knowledge of both tangent lines and roots is required for students to understand and apply Newton's method. Analysis was done from the perspective of the empirical, implied, and intended readers and was examined for the extent to which the empirical and implied readers aligned. It was found that although the alignment of the empirical and implied readers was helpful in determining the success of the students, it was not in itself a deciding factor.

Keywords: Newton's method, Reader oriented theory, Calculus, Tangent lines

Newton's method is an iterative technique using tangent lines and x-intercepts to find the root of a function. Because of the distinct concepts involved and where it falls in the curriculum, Newton's method provides an ideal and yet unexplored case study for reader oriented theory. Newton's method is an opportunity for students to put their newly acquired knowledge of tangent lines to use while integrating it with previous knowledge of x-intercepts and roots. This study examines the prerequisite knowledge required to correctly apply Newton's method and the capabilities of students who lack some or all of this material. Students were asked to read a paragraph describing how Newton's method works and were presented with a graph and initial guess. They were then asked to follow the steps presented to implement Newton's method. We find that alignment of implied and empirical readers is an informative but imperfect predictor of success.

Literature Review

An understanding of the concepts of derivative and of tangent line is important for interpreting Newton's method correctly. While drawing the line tangent to a curve, it may be helpful for students to interpret the derivative as the slope of the tangent line. Newton's method requires that students draw the line tangent to a curve at a chosen point and follow it to where it crosses the x-axis. Newton's method requires the student to draw on the graphical, verbal, and symbolic representations of derivative as presented in Zandieh's (2000) derivative framework.

The concept of the derivative has been well researched. "It is known that some students are introduced to differentiation as a rule to be applied without much attempt to reveal the reasons for and justifications of the procedure." (Orton, 1983, p. 242) In fact, many first semester calculus students survive without ever developing a conceptual understanding of the derivative. Asiala, Cottrill, Dubinsky, and Schwingendorf (1997) found evidence that students strongly feel the need for an expression of a function they can differentiate, rather than being able to interpret the derivative as the slope of the tangent line at a point from a graph. They found that students sometimes think the derivative and the equation of the tangent line at a point are the same thing.

Students' understanding of drawing tangent lines has been studied to a lesser extent than has the concept of derivative. Many have identified students' struggles with drawing lines tangent to curves (Biza, 2011; Biza, Christou, & Zachariades, 2008; Kajander & Lovric, 2009; Páez Murillo & Vivier, 2013). Biza et al. (2008) found that students have difficulty drawing tangent lines at inflection points, cusps and at edge points (discontinuities created by breaks in the graph). These misconceptions may have been perpetuated by textbooks and teachers alike. Mathematics textbooks have said that a "tangent line touches the graph of $y = f(x)$ at only one point." The word touching implies that the line does not share anything in common with the curve, while in fact they do share a point in common. Diagrams in textbooks tend to show a tangent line that only passes through the curve at one point and do not consider special cases, such as cusps, inflection points, and edge points (Kajander & Lovric, 2009).

Theoretical Perspective

Our theoretical framework comes from reader oriented theory (Weinberg & Wiesner, 2011). According to reader oriented theory, there are three important terms: the intended reader, the implied reader, and the empirical reader. The *intended reader* is the particular audience that the author expects to read the text. This can be seen in the language that the text uses. For example, the use of "we" in mathematical texts when addressing the audience implies that the reader is a member of the mathematical community. The *implied reader* is the set of skills the reader will need in order to understand the text. For example, the implied reader needs to carry out certain behaviors indicated in the text and should have certain mathematical skills and knowledge. Of particular importance to understanding Newton's method is the definition of tangent line and root and the ability to draw a tangent line. The *empirical reader* is the actual person who reads the text. The empirical reader's concept images will impact how they interpret the text. A concept image is "the set of all the mental pictures associated in the student's mind with the concept name, together with all the properties characterizing them" (Vinner & Dreyfus, 1989, p. 356). A student's concept image does not necessarily correspond to the mathematical definition of the term. In some cases, students' responses are shaped by their desire to please an authority figure. When this happens they may demonstrate pseudo-conceptual behavior, giving answers that mirror cognitive thought but are not actually derived from cognitive thought (Vinner, 1997).

We expected that students who had seen Newton's method previously would demonstrate a higher level of reading comprehension. There is research indicating that children who have prior knowledge of a topic show greater fluency and word recognition when reading ability is held constant (Priebe, Keenan, & Miller, 2012). We are curious to know if this also holds for older students.

Methods

We created a text for Newton's method. The aim was to identify characteristics of the implied and empirical reader and how those characteristics fostered comprehension of the text. Interviews were conducted on a voluntary basis with twelve students enrolled in first semester calculus at a large university. Six of the participants had previously taken calculus courses, and three of these students had seen Newton's method prior to the interview. Interviews were video recorded and all written work was kept for analysis.

Interviews were completed prior to instruction on Newton's method in the classroom, but subsequent to instruction on derivative, tangent line, and the Intermediate Value Theorem (another tool for approximating roots). Prior to reading, students were asked to give a definition of tangent line as well as examples and non-examples. Each student was asked to silently read a brief passage describing the uses of Newton's method, which was adapted by the research team from three calculus textbooks. Following the background passage, the student silently read a paragraph describing how to use Newton's method. After completing the reading, the student was asked to re-read the final paragraph aloud and ask any questions he or she may have. The student was then given the chance to demonstrate his or her understanding of the text by carrying out Newton's method graphically on the example provided by being asked to find x_2. This was followed with additional questions pertaining to Newton's method and the reading.

Upon completion, the interviews were transcribed by the authors for analysis. Though initial analysis presented several categories of interest, three main ideas were selected for discussion here: the tangent line concept, the root concept, and whether the student was able to follow the reading to demonstrate Newton's method. For the tangent line concept, the students were partitioned into three categories based on their understanding: conceptual, pseudo-conceptual, and none. Those with a conceptual understanding drew various tangent lines correctly and used key phrases relating to the correct definition. Those with a pseudo-conceptual understanding could draw some tangent lines, but struggled to do so at non-extreme values. With regard to the root concept, students were again partitioned into three categories. Students with prior knowledge were able to give the definition of root prior to performing Newton's method while students with a developed knowledge used Newton's method to help them articulate a definition and students with no knowledge were unable to give a definition at any point during the interview. Finally, the students were once more partitioned into three categories based on their ability to follow Newton's method. Students who were able to follow an iterative process that mimicked Newton's method but with incorrect tangent lines or using x-values other than the x-intercept were categorized as having a partial ability. The remaining students were either fully able or unable to follow Newton's method.

Results

The implied reader of the Newton's method document has an understanding of tangent lines and roots. These are concepts that first semester calculus students should have encountered by this point in the semester. In the following sections, the interviews are categorized based on the students' understanding of the concepts of tangent line and root. Analysis of the interviews focused on these concepts to determine whether each was pivotal in the ability to carry out Newton's method based on the reading. Table 1 presents the subjects and categorizes their knowledge of these two primary subjects along with their ability to carry out Newton's method.

Student	Tangent Line Concept	Root Concept	Ability to Follow
Aaron	Pseudo-conceptual	Developed	Partial
Andrea	Pseudo-conceptual	None	No
Bethany	Pseudo-conceptual	None	No
Dave	Conceptual	Developed	Yes
Janet	Conceptual	Prior	Yes
Kevin	Conceptual	Prior	Yes
Larry	None	Developed	Yes
Lewis	Pseudo-conceptual	Prior	Partial
Mark	Pseudo-conceptual	None	Yes
Qadan	Pseudo-conceptual	None	Partial
Steve	Conceptual	None	No
Ted	Conceptual	Developed	Yes

Table 1: Summary of prerequisite concepts and the ability to follow Newton's method.

Tangent Line Concept

The students interviewed were categorized into three groups based on their understanding of tangent lines: with a conceptual understanding, with a pseudo-conceptual understanding, and with no understanding. The level of understanding indicates the alignment of the empirical and implied readers with respect to the tangent line concept.

Conceptual understanding. Five students had a conceptual understanding of tangent lines, meaning the empirical and implied readers were fully aligned. Although none of these students could effectively verbalize the definition of a tangent, some used key related phrases such as:

Kevin: The tangent line is a representation of the instantaneous rate of change . . . It is the derivative. But it's the slope of what you get the derivative from

Dave: It's a line that, uh, passes through a point on a curve once.

Despite their struggles verbally, each of these students was able to draw tangent lines correctly regardless of the shape of the function and the chosen x value. Each student had a robust concept image relating to tangent lines. One of these students, however, was unable to apply Newton's method correctly.

Pseudo-conceptual understanding. Six of the students showed some knowledge of tangent lines but did not demonstrate a complete understanding. This means the empirical and implied readers overlapped, but did not fully align. These students showed a pseudo-conceptual understanding of tangent lines, generally focusing only on horizontal tangent lines.

Andrea is very clear at the beginning of the interview that she believes a tangent line only occurs where there is an extreme value. When asked to draw tangent lines on a graph, she draws tangent lines at the maximum and minimum, saying "right there would be a tangent line because the slope would be zero there." When asked to draw a tangent line at a point on the graph with a non-zero derivative, she said "because the slope is constantly changing there, so you could not draw a tangent line." This caused difficulties in Andrea's ability to carry out Newton's method because a non-horizontal tangent is required in order to approximate the root appropriately. This pseudo-conceptual understanding could not be remedied by the interviewer drawing connections to tangent lines and circles. Andrea was unable to perform Newton's method despite being given the definition of tangent line during the interview.

Bethany was not able to successfully carry out the Newton's method process because her pseudo-conceptual understanding of tangent lines was a significant obstacle in her taking meaning from the passage. In drawing tangent lines, she started with a horizontal tangent line and then tilted it until it became a secant line (Figure 1). After the reading, she tried to draw a tangent line that followed the curve exactly so that her line would intersect the x-axis by tilting the tangent line from the horizontal one.

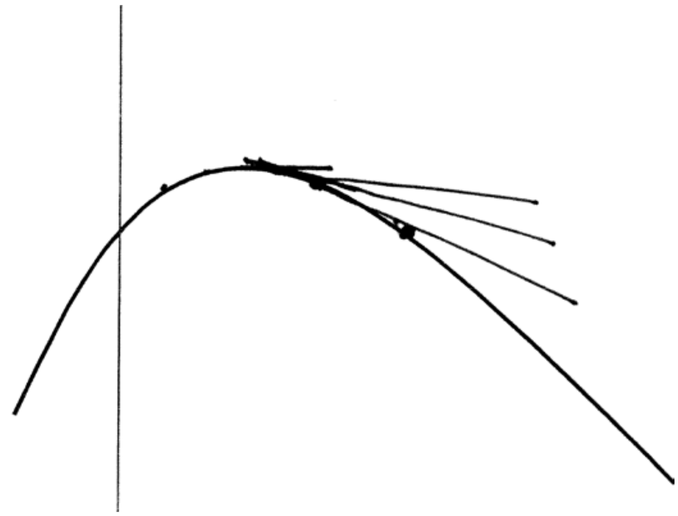


Figure 1: Bethany's process of drawing tangent lines via tilting.

This "tilting" process was also modeled by Aaron, yet in a slightly different manner. Instead of beginning with a tangent line and morphing it into a secant line, he used a tilting argument to show how the slope of the tangent line changes. When asked how he determined the slope of the tangent line, he explained "I just kinda take the middle point of it and start moving it down and slowly I can imagine in my mind like tipping over a little bit" (Figure 2). Despite his incorrect tangent lines, Aaron was able to correctly follow the reading to perform a series of steps that created a succession of x values getting closer to the root. Hence, he seemed to understand the basic idea of what the described iterative process was intended to achieve.

Qadan drew tangent lines similar to Aaron's. Although he was able to determine the sign of the slope of the tangent line and how it related to the slopes of the tangent lines at surrounding points, he was unable to draw the tangent line correctly. Like Aaron, he was successful in performing an iterative process of finding successive x values.

Mark and Lewis were both on the verge of a conceptual understanding of tangent lines. Mark knew that the tangent line should only pass through one point on the curve, but he didn't realize that this only means locally. He was able to remedy his thinking and perform Newton's method correctly later in the interview. Lewis was able to draw several correct tangent lines at the beginning of the interview, but drew them through the graph when applied to Newton's method. He continued to talk about lines through the curve being inaccurate, saying "that was just an example, like it would be more like on, towards how the actual graph is like going." Though he did not perform Newton's method correctly, he was successful in performing an iterative set of steps based on the reading.

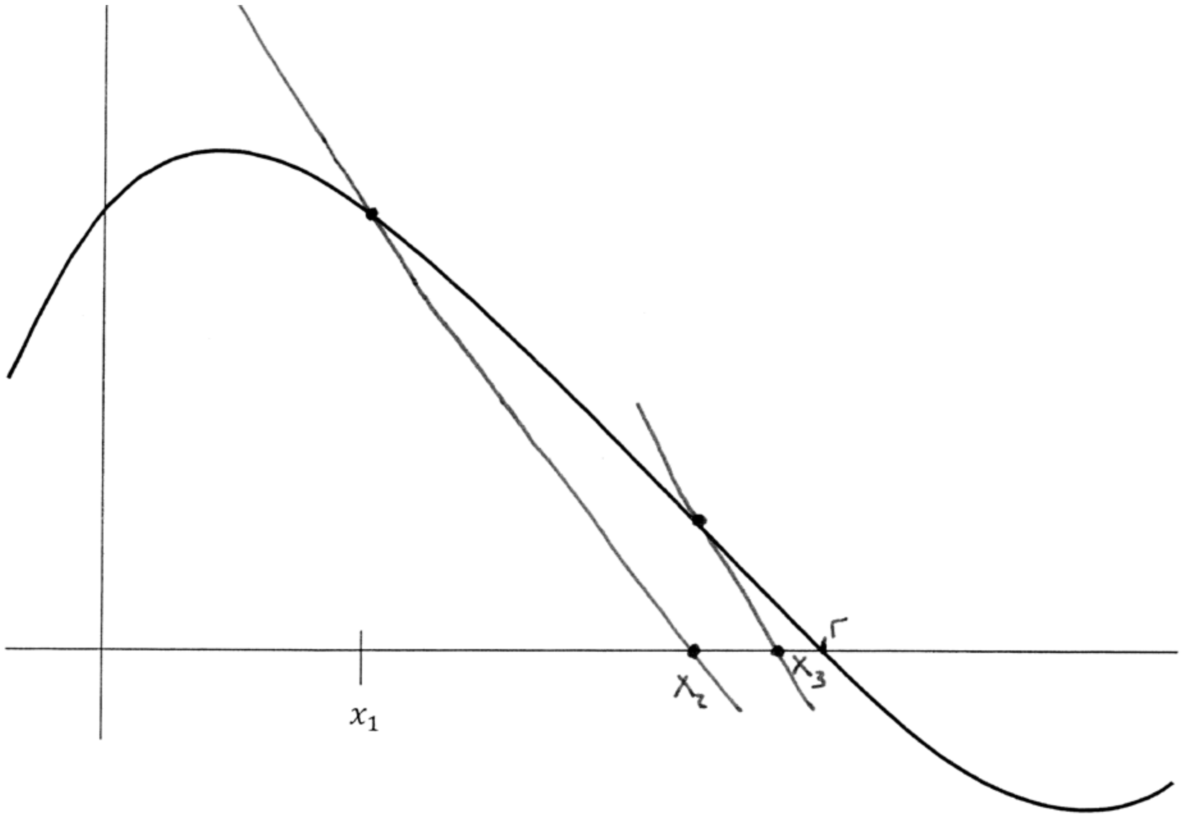


Figure 2: Aaron performs an iterative process with incorrect tangent lines.

No understanding. Larry had no understanding of tangent lines. The empirical and implied readers did not align at all. He began to try and describe a tangent line as "a line of an angle at [pause] a horizontal point," but abandoned this idea. His next attempt was at drawing the graph of the trigonometric function tangent. He ultimately decided that he did not know what a tangent line is. The definition and a few examples were then given to him by the interviewer. After receiving this information, he was able to carry out Newton's method correctly.

The Meaning of Root

The students were categorized into three groups based on their understanding of root. The empirical reader and implied reader were aligned prior to working through the Newton's method process, they became aligned while working through the process, or they did not align at all.

Prior understanding. Three students had a clear conceptual understanding of root prior to carrying out Newton's method. Two of these students were clear from the beginning, while one had to draw on past knowledge to recall the meaning of root. Two of these students were able to carry out Newton's method correctly, while the third was able to construct an iterative process based on the reading. Understanding what was being approximated helped guide the process being performed and clarify the end goal.

Developed understanding. Four students had no understanding of root prior to working through Newton's method but were able to deduce the meaning while working through the reading on the Newton's method process. Dave initially placed r arbitrarily on the x-axis and did not relate it to the root of the function. After working through the process to find x_4, he saw that it was farther away from what he marked as r and recognized that there was a problem. In realizing what he was trying to approximate, he was able to determine the meaning of root.

Int 2: So what should your r be?
Dave: Or uh it's close to x.
Int 2: What's the root of a function generally speaking?
Dave: Um where the function equals zero.
Int 2: Ok, so where does that function equal zero?
Dave: Oh right there. Ohhhhh look at that.

These students were able to use the given context to connect the root with the value being approximated by Newton's Method, thereby aligning the implied and empirical readers.

No understanding. Five students had no understanding of root prior to working through Newton's method and were unable to use Newton's method to deduce the meaning of root. One of these students actually knew the meaning of root, but was considered to have no understanding because she was not confident in her definition saying "I'm calling this [points to the root] like the, the root. Which I don't think is exactly right, but that's what I've been doing." Of these five students, one was able to carry out Newton's Method successfully, one was able to perform an iterative process based on the reading, and the other three were unsuccessful.

Discussion

The implied and empirical readers should align with respect to tangent lines to facilitate the students' ability to take what was read and turn it into a mathematical process. Although some students were able to carry out an iterative process that produced subsequent x-values without a proper understanding of the tangent line concept, the final graphic was not a true representation of Newton's method. Students frequently had a pseudo-conceptual understanding of tangent lines consisting primarily of horizontal tangent lines, likely due to this being the most recent application of tangent lines covered in class. This led to difficulties for students as the tangent lines required to perform the task must not be horizontal. One thing an instructor may do to increase the success of the student is to be sure to present a range of tangent lines and help students construct several of their own non-horizontal tangent lines in order to build a solid conceptual understanding. Engaging the students in constructing many examples and non-examples helps to create a well-connected example space from which they can later draw when solving new and novel problems. (Sinclair, Watson, Zazkis, & Mason, 2011) We expect that a robust example space from which to draw when reading a text would increase success rates.

To a lesser extent, the implied and empirical readers should align with respect to the root concept. Understanding what they were trying to approximate greatly benefited the students who could not correctly follow the process at first. This knowledge facilitated the student's ability to see how to draw a tangent line when the function was below the x-axis. For other students, working through the reading helped develop their concept of root. Hence, even though the implied and empirical readers were not initially aligned, the text provided an opportunity to develop the required root concept, align the readers, and encourage success.

Possessing knowledge of both tangent lines and roots still may not result in success. A common trait among the three students who were unable to translate the reading into a mathematical process was a lack of confidence. Two of the students made several comments relating to their mathematical struggles and made little effort to really make sense of the material. The third student's lack of confidence created enormous confusion while she was attempting to interpret the reading. Building student confidence by incorporating mathematical reading exercises into the curriculum may lead to increased understanding.

References

Asiala, M., Cottrill, J., Dubinsky, E., & Schwingendorf, K. (1997). The Development of Students' Graphical Understanding of the Derivative. *Journal of Mathematical Behavior, 16*(4), 399-431.

Biza, I. (2011). Students' evolving meaning about tangent line with the mediation of a dynamic geometry environment and an instructional example space. *Technology, Knowledge and Learning, 16*(2), 125-151. doi: 10.1007/s10758-011-9180-3

Biza, I., Christou, C., & Zachariades, T. . (2008). Student perspectives on the relationship between a curve and its tangent in the transition from Euclidean geometry to analysis. *Research in Mathematics Education, 10*(1), 53-70. doi: 10.1080/14794800801916457

Kajander, A., & Lovric, M. (2009). Mathematics textbooks and their potential role in supporting misconceptions. *International Journal of Mathematical Education in Science and Technology, 40*(2), 173-181. doi: 10.1080/00207390701691558

Orton, A. (1983). Students' Understanding of Differentiation. *Educational Studies in Mathematics, 14*, 235-250.

Páez Murillo, R. E. , & Vivier, L. (2013). Teachers' conceptions of tangent line. *Journal of Mathematical Behavior, 32*(2), 209-229. doi: 10.1016/j.jmathb.2013.02.005

Priebe, S. J., Keenan, J. M., & Miller, A. C. . (2012). How prior knowledge affects word identification and comprehension. *Reading and Writing, 25*(1), 131-149. doi: 10.1007/s11145-010-9260-0

Sinclair, Nathalie, Watson, Anne, Zazkis, Rina, & Mason, John. (2011). The structuring of personal example spaces. *The Journal of Mathematical Behavior, 30*, 291-303. doi: 10.1016/j.jmathb.2011.04.001

Vinner, Shlomo. (1997). The Pseudo-Conceptual and the Pseudo-Analytical Thought Processes in Mathematics Learning. *Research Studies in Mathematics, 34*, 97-129.

Vinner, Shlomo, & Dreyfus, Tommy. (1989). Images and Definitions for the Concept of Function. *Journal for Research in Mathematics Education, 20*, 356-366.

Weinberg, Aaron, & Wiesner, Emilie. (2011). Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics, 76*(1), 49-63. doi: 10.1007/s10649-010-9264-3

Zandieh, Michelle. (2000). A Theoretical Framework for Analyzing Student Understanding of the Concept of Derivative. In E. Dubinsky, Schoenfeld, A., Kaput, J. (Ed.), *Research in Collegiate Mathematics Education, IV* (Vol. 8, pp. 103-127). Providence, RI: American Mathematical Society.

AN EXPLORATION OF MATHEMATICS GRADUATE TEACHING ASSISTANTS' TEACHING PHILOSOPHIES

Kedar Nepal
Oklahoma State University

This is an investigation of the teaching philosophies of beginning mathematics graduate teaching assistants. Three teaching philosophy statements from each of four participants were collected at different stages of a semester-long teaching assistant preparation program and analyzed. Principal elements found in these statements before they underwent training and how their philosophies changed over time during training will be discussed.

Key words: Teaching Philosophy, Mathematics Graduate Teaching Assistants, Undergraduate Mathematics Instruction, GTA Training.

Introduction and Research Questions

Studies have shown that teachers' existing knowledge and beliefs inform and guide their classroom practices and decisions (Kim, 2011; Speer, 2008; Thompson, 1992). Since teachers' teaching philosophies (collection of existing beliefs) change over time, careful examination of beginning and evolving teaching philosophies may provide insights into the support structures necessary to facilitate effective classroom instruction (Simmons et al., 1999). However, very few graduate teaching assistant (GTA) preparation programs offered by mathematics departments have attempted to understand and incorporate teachers' beliefs and perspectives in their preparation programs. Instead, most programs are designed based on the faculty's wisdom and experience (Chae, Lim, & Fisher, 2009). Because most math GTAs ultimately adopt a career teaching collegiate mathematics, it is critical that their professional development be studied and nurtured. However, little is known about beginning GTAs' teaching philosophies, how their philosophies change during their pre-service and in-service phases, and what factors affect GTAs and their teaching philosophies. This qualitative study therefore attempts to examine the evolution of GTAs' teaching philosophies, both as pre-service and in-service instructors.

Research questions

1. What are the teaching philosophies of beginning mathematics GTAs?
2. How do their philosophies evolve during the pre-service phase?
3. How are their evolving philosophies nurtured, and how do they change as they transition to an in-service phase?
4. What are the major contributing factors that affect mathematics GTAs and their teaching philosophies during the pre-service and in-service phases?

Methods

This study was guided by the context-based adult learning (Hansman, 2001), an extension of Vygotsky's sociocultural theory to adult learning. An assumption to this study was that GTAs' beliefs are developed, changed or reinforced as they learn more about teaching and learning, and these changes are reflected in their teaching philosophies.

In the Fall 2012 semester, the researcher observed all the class sessions of a semester-long GTA preparation program course (see Appendix A) offered by the mathematics department at a large public university in the Midwestern US. This was a mandatory pre-service program for all beginning mathematics GTAs, who did not need to teach or assume any other TA duties during that semester. In order to collect data from the participants, the instructor of the course required them to submit three teaching philosophy statements (TPS) in three installments: TPS I at the beginning, TPS II in the middle, and TPS III at the end of the semester. The prompts for writing these TPSs were discussed between the instructor and the researcher before they were given to the GTAs (see Appendix C). After the completion of the course, a purposeful sampling method was employed to select four of those GTAs. In order to include the maximum variation possible, two of the four GTAs selected were domestic, the other two were international students and each category included a female and a male participant. All of these GTAs are currently assuming their TA duties. See Appendix B for pseudonymous descriptions of the four GTAs.

The participants' teaching philosophy statements were collected and analyzed using open coding techniques. The researcher also conducted one-on-one interviews with each participant during the Summer 2012 and the Fall 2013 semesters, which were audio-taped for transcription and analysis. One more one-on-one interview is scheduled for the Spring 2014 semester.

Preliminary Results

All teaching philosophy statements have been coded, but the interviews are still being transcribed and analyzed. Therefore data pertaining only to research questions 1, 2, and part of 4 are currently available.

Below is a list of the principal themes identified in each of the GTAs' teaching philosophy statements, and also the factors that influenced their philosophies. I expect more refined themes to emerge once all the data have been collected and analyzed. Due to space constraints, I was not able to elaborate on either of these, or to provide sample quotes from the GTAs. Readers may find it difficult to understand the results without further elaboration of these themes and factors; however, I hope to provide detail description of each of these themes and factors in future publications.

David. TPS I: positive attitude, concept, content knowledge, equal treatment, high expectation for students, sense of humor, and creation of positive feeling towards mathematics. TPS II: all themes from TPS I except content knowledge, plus organization, personality, preparation, clarity of speech and tone, interaction, promotion of learning, caring, concept as well as process. TPS III: all themes from TPS II, plus preparation, experience, carefulness, classroom management, dealing with students, friendliness, language fluency, cultural understanding.

Factors: High school and undergraduate experience, a high school teacher, classroom practicum, language, culture, undergraduate students.

Andrew. TPS I: high expectations for students, being tough with students, application of knowledge, engaging classroom environment, preparation, balance between teaching and other duties, out-of-class support, coping with institutional culture. TPS II: all themes from TPS I, plus interaction. TPS III: all themes from TPS II, plus inspiration and encouragement, positive attitude, making students think, concept as well as process, self-reflection, solving problems using multiple techniques, promotion of individual development.

Factors: a math professor in undergraduate program, institutional context, personal needs, classroom practicum, undergraduate students, subjects learned other than mathematics.

Rebecca. TPS I: passion for mathematics and teaching, inspiration and encouragement, caring, trust, high expectation, concept, enthusiasm, engaging classroom environment, promoting collaboration, out-of-class support, application of knowledge. TPS II: all themes from TPS I, plus organization, preparation, attitude. TPS III: all themes from TPS II, plus a welcoming classroom environment.

Factors: undergraduate professor, past experience learning and teaching, classroom observation, mentor.

Jennifer. TPS I: content knowledge, positive attitude, friendliness, equal treatment, caring, out of class support, respecting students' opinions and ideas, inspiration and encouragement, making students independent learners, application of knowledge. TPS II: all themes from TPS I, plus eye contact. TPS III: all themes from TPS II, plus a welcoming classroom environment, patience, respect for students' privacy, acknowledgement of differing capabilities among students.

Factors: personal background, high school and undergraduate experience, a high school teacher, subjects learned other than mathematics, classroom observation, classroom practicum, teaching by faculty in own coursework.

As expected, GTAs expressed varying ideas and beliefs about teaching and learning. However, some elements in their TPSs were common to most of the GTAs. Also, some elements were found common to a specific group or category (e.g., females, international GTAs). In their TPS I, all but Jennifer expressed that having high expectations of students (such as giving hard and challenging problems in homework and exams) could force them to think and work hard. All but Andrew wrote that a teacher should have a positive attitude towards teaching. Similarly, most GTAs believed that instructor should relate mathematical concepts taught in the classroom to real life problems, in order to motivate and prepare students to apply their knowledge. All GTAs except David underscored the importance of encouraging or inspiring students and helping them outside of classroom.

It was interesting to observe that the international GTAs, David and Jennifer, believed that teachers should treat their students equally and that teachers' content knowledge was key to their success. Both domestic students, Andrew and Rebecca, believed that instructors should motivate students to think, learn and succeed, instead of just transferring their own knowledge to the students. They also expressed that teachers should keep their students engaged in the classroom.

The female GTAs expressed that teachers should have a caring attitude: they should care about their students' success. Jennifer believed that students should not be judged by their exam grades. Rebecca believed that a teacher should win students' trust. Rebecca was the only GTA to express anything about student collaboration, something that most educators think is beneficial to student learning. Rebecca also believed that students should see their teacher's enthusiasm and passion towards the subject. Andrew believed (in TPS I) that a teacher should maintain a balance between his instructional duties and personal life. He also emphasized the need to cope with the institutional culture. He believed that teachers should employ tough love attitude with the students. According to him, being 'too nice' with students does not help them succeed.

No GTAs stated that their earlier opinions from TPS I had changed. Instead, they all repeated the opinions they had expressed in TPS I, but expressed additional opinions in their later statements. In TPS II and III, David expressed that content knowledge alone is insufficient, and that a teacher needs to be skilled in preparation, interaction, organization, teaching techniques, and speaking clearly. In his TPS III, he also expressed that teachers should be fluent in English and have a strong understanding of the American culture. On the other hand, Jennifer emphasized teachers' content knowledge in all her TPSs, and never mentioned the importance of fluency in English or cultural understanding. It is interesting to note, however, that David was much more fluent in English than Jennifer. Also, in TPSs II and III, both of them repeated their earlier position that teachers should treat their students equally. It will be interesting to see how these GTAs' philosophies evolve during their in-service phase.

All the GTAs were influenced more than anything by the teaching they had experienced during their undergraduate or high school times, especially by the role model teachers they had. Pre-service classroom practicum also had some influence on their teaching philosophies.

I am considering asking one or two peers to code TPSs of at least one GTA to establish intercoder reliability. My question to the audience: What would be a better way to elicit their current teaching philosophies? How can we detect changes in their philosophies? What could be done to corroborate findings?

References

Belnap, J. (2005). *Putting TAs into context: Understanding the graduate mathematics teaching assistant.* Retrieved from *ProQuest Dissertations and Theses.* (305024286)

Bonk, C. J., & Kim, K. A. (1998). Extending sociocultural theory to adult learning. In M. C. Smith & T. Pourchot (Eds.), *Adult Learning and Development: Perspectives From Educational Psychology* (pp. 67-88). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brandt, B. L., Farmer, J. A., & Buckmaster, A. (1993). Cognitive apprenticeship approach to helping adults learn. *New Directions for Adult and Continuing Education, 1993*(59), 69-78.

Chae, J., Lim, J. H., & Fisher, M. H. (2009). Teaching mathematics at the college level: International TAs' transitional experiences. *Primus, 19*(3), 245-259.

Chism, N. V. N. (1998). Developing a philosophy of teaching statement. *Essays on Teaching Excellence: Toward the Best in the Academy, 9,* 1–3.Teaching mathematics at the college level: International TAs' transitional experiences. *Primus, 19*(3), 245-259.

Brandt, B. L, Farmer, J. A., & Buckmaster, A. (1993). Cognitive apprenticeship approach to helping adults learn. *New Directions for Adult and Continuing Education*, 59, 69-78.

Hansman, C. A. (2001). Context-based adult learning. *New Directions for Adult and Continuing Education,* 89, 43-51.

Harris, G., Froman, J., F., & Surles, J. (2009). The professional development of graduate mathematics teaching assistants. *International Journal of Mathematical Education in Science & Technology, 40*(1), 157-172. doi: 10.1080/00207390802514493

Jenkins, C. (2011). Authenticity through reflexivity: Connecting teaching philosophy and practice. *Australian Journal of Adult Learning, 51,* 72-89.

Jones, J. L. (1993). TA training: From the TA's point of view. *Innovative Higher Education, 18, 147-161.*

Kim, M. (2011). *Differences in beliefs and teaching practices between international and U.S. domestic mathematics teaching assistants.* Retrieved from ProQuest Dissertations and Theses. (885228899)

Kung, D., & Speer, N. (2007). Mathematics teaching assistants learning to teach: Recasting early teaching experiences as rich learning opportunities. *Electronic Proceedings for the Tenth Special Interest Group of the Mathematical Association of America on Research on Undergraduate Mathematics Education.*

Simmons, P. E., Emory, A., Carter, T., Coker, T., Finnegan, B., Crockett, D., Labuda, K. (1999). Beginning teachers: Beliefs and classroom actions. *Journal of Research in Science Teaching, 36*(8), 930-54.

Speer, N. M. (2008). Connecting Beliefs and Practices: A Fine-Grained Analysis of a College Mathematics Teacher's Collections of Beliefs and Their Relationship to His Instructional Practices. *Cognition and Instruction,* 26(2), 218-267.

Speer, N., Gutmann, T., Murphy,T. J. (2005). Mathematics teaching assistant preparation and development. *College Teaching*, *53*(2), 75-80.

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York: Macmillan.

Vygotsky, L. S. (1978). *Mind in the Society: The Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Appendix A

Description of GTA Preparation Course

This is a pre-service preparation program offered every fall semester by the mathematics department to train beginning GTAs. They do not need to teach or assume any other TA duties during their first semester in the graduate program. GTAs in this course learn from weekly seminars and a classroom practicum.

Classroom practicum. GTAs are placed with experienced instructors who serve as mentors. GTAs are expected to observe their mentors' class sessions and participate in all activities assigned by them. GTAs are also expected to maintain logs of their practicum experiences in their course portfolios. Following each class meeting, they write down their observations, questions, and reflections in their practicum logs. Periodically, the logs are reviewed by the instructor. GTAs prepare and deliver a few actual classroom presentations under the direct supervision of their mentors. GTAs write reflections on their own presentations and discuss these reflections with peer GTAs, who also write reflective comments related to their observations of the presenter. Mentors submit an evaluation of the GTAs' performance to the course instructor at the end of the semester.

Weekly seminars. GTAs are expected to complete all out of class assignments, such as writing syllabi, lessons, exams, and papers, and retain them in their course portfolios. Most seminars begin with a discussion of typical decision making and classroom management issues related to undergraduate education. Topics included make-up requests, cheating, responding to student emails. As the class progresses, the participants are asked to share their observations, questions, and reflections they have noted from observing their mentors' classroom. GTAs are asked to learn routine activities such as preparing syllabi, writing quizzes and exams, using technology in the classroom, maintaining a grade book, and posting student grades from their mentors. They routinely share their observations with other GTAs and the course instructor.

GTAs are also required to grade actual student homework and exams. After doing so, they display their grades and grading algorithm on the board and are asked to justify their decision. Besides several other reading and writing assignments, GTAs are assigned a particularly introspective assignment related to 'Developing Your Philosophy of Teaching'.

After the completion of this pre-service program, they enter into the in-service phase. They begin to teach and assume other TA related duties such as grading and tutoring during this phase.

Appendix B

Description of participants

David (age 24) is an international graduate student from a south Asian country. He finished his undergraduate degree from a medium sized university in the midwestern United States. He completed high school in his home country. His only teaching experience was tutoring mathematics at his undergraduate institution. He conducted two recitation sections of business calculus course in the Spring 2013 semester, and did not teach any course in the Summer 2013. He is a master's student and is willing to pursue a Ph.D. degree in applied mathematics. He wants to become a professor after finishing his degree. He is fluent in English with a foreign accent.

Andrew (27) is a domestic graduate student who completed school in the southwestern United States, and finished his undergraduate degree from a university in the same region. He was home schooled during his high school period. He is a Ph.D. student and wants to work in industry. His only teaching experience was tutoring undergraduate students. He taught two sections of the Functions in Spring 2013, and a section of Business Calculus during Summer 2013.

Rebecca (23) is a domestic student who completed high school in the midwestern US, and received her undergraduate degree from a small Catholic university in the same region. Her only teaching experience before joining her graduate program was tutoring undergraduate students. She taught one section of the Pre-calculus during Spring 2013 and a section of College Algebra during Summer 2013. She is a master's student and does not have any plan to pursue her Ph.D. She wants to go to a business field to work in the future.

Jennifer (30) is an international student who completed her high school and college education up to masters' degree from a north-eastern Asian country. She taught mathematics in a college for five years in her home country. She taught two recitation sections of Business Calculus in Spring 2013, but did not teach any course during Summer 2013. She is a Ph.D. student and wants to become a mathematics professor in the future. She speaks English with a foreign accent but is not as fluent as David.

Appendix C

Prompts for writing teaching philosophy

TPS I: Write a short paper of about 2-3 pages discussing what *you* have learned about effective and ineffective teaching from being a student. Describe the teaching of someone who was, in your experience, a particularly effective teacher, and analyze why you think this person succeeded as a teacher. This is just the beginning on your journey to develop your own philosophy of teaching, a philosophy that will probably change several times during your teaching career. The conclusion of your paper should be a thoughtful initial statement of *your* emerging philosophy of teaching. Be sure to include your thoughts on what you believe now.

TPS II: Earlier this semester you made an initial effort to characterize your philosophy of teaching. In this paper, you discussed what you have learned about effective and ineffective teaching from being a student. You described the teaching of someone who was, in your experience, a particularly effective teacher, and analyzed why you thought this person succeeded as a teacher. You concluded this paper with an emerging statement of your philosophy of teaching. Your job now is to revisit this paper, rethink its contents, and revise it based on reflections on your practicum and seminar experiences this semester - so far. You will have yet another opportunity to reflect as we approach the end of the semester.

TPS III: Earlier this semester you made efforts to characterize your philosophy of teaching. In your first paper, you discussed what you had learned about effective and ineffective teaching largely from the perspective of being a student. You described the teaching of someone who was, in your experience, a particularly effective teacher, and analyzed why you thought this person succeeded as a teacher. And, you concluded this paper with an emerging statement of your philosophy of teaching. You revisited this paper about midway through the semester, rethought its contents, and revised it based on reflections of your practicum and seminar experiences this semester - so far. Now, you have the chance to revise it yet a final time. This final revision provides you with an opportunity to build into your philosophy statement those most recent experiences in the classroom as a teacher. Hopefully, you will recognize and appreciate the journey we have taken this semester - a journey that began with your beliefs about teaching being based on your experiences as a student and is ending with your beliefs about teaching being based on learning first-hand what is involved in the art of teaching.

Preservice Elementary Teachers' Understanding of Number Theory: Connecting Content Knowledge to PCK

Many preservice elementary teachers have a limited understanding of the mathematics that they will teach, including many topics in number theory (e.g., Zazkis & Liljedahl, 2004), which suggests that they may not be prepared to teach mathematics for understanding. The research also suggests that pedagogical content knowledge (PCK) is important for teaching (e.g., Ball, Thames, & Phelps, 2008; Shulman, 1986), but little is known about preservice elementary teachers' PCK in number theory.

The overarching research question of this study is: What is the nature of preservice elementary teachers' understanding of topics in number theory? The primary goal was to investigate preservice elementary teachers' content knowledge and PCK in number theory, with topics such as greatest common factor (GCF), least common multiple (LCM), and prime numbers. I also explored opportunities preservice elementary teachers have to develop number theory PCK as well as if and how their number theory content knowledge might contribute to their number theory PCK. I share my findings from the latter endeavor in this report.

Background

Number theory content is integrated throughout primary school mathematics education, ranging from even and odds in pre-Kindergarten to prime factorization, GCF, and LCM in the middle grades. Zazkis and colleagues contributed the bulk of what little is known about preservice elementary teachers' understanding of topics in number theory, such as even and odd numbers (Zazkis, 1998), multiplicative structure (Zazkis & Campbell, 1996), and prime numbers (Zazkis & Liljedahl, 2004). In general, most participants exhibited a procedural understanding of the content and difficulty working flexibly with various number concepts. For instance, Zazkis (1998) found that participants struggled to associate "evenness" with "divisibility by 2" and having a factor of 2 in the prime factorization. Brown, Thomas, and Tolias (2004) investigated preservice elementary teachers' understanding of LCM and found that many participants had a similarly procedural understanding of the content.

While it is unclear how a teacher's content knowledge may affect student learning, the research suggests that a teacher's PCK does impact teacher effectiveness (Shulman, 1986). Ball and colleagues have further conceptualized mathematical PCK by proposing subconstructs such a knowledge of content and students (KCS), knowledge of content and teachers (KCT), and knowledge of curriculum. According to Hill, Schilling, & Ball (2004), KCS pertains to "knowledge of students and their ways of thinking about mathematics – typical errors, reasons for those errors, developmental sequences, strategies for solving problems". KCT requires "coordination between the mathematics at stake and the instructional options and purposes at play" (Ball, Thames, & Phelps, 2008). Ball, Thames, and Phelps also define a type of content knowledge specific to teachers called specialized content knowledge (SCK), defined as "the mathematical knowledge that allows teachers to engage in particular teaching tasks, including how to accurately represent mathematical ideas, provide mathematical explanations for common rules and procedures, and examine and understand unusual solution methods to problems" (p. 377-8). Little is known about preservice elementary teachers' number theory PCK or SCK.

The emergent perspective (Cobb & Yackel, 1996) served as the lens for collecting and analyzing data. I primarily used the psychological lens since the bulk of the data represent individual conceptions about number theory content. On the other hand, via the social lens I explored the classroom norms, expectations, and experiences that framed participants' perspectives on number theory content and how they might use it to teach. I also drew from Ball and colleagues' conceptualization of mathematical PCK (e.g., Ball, Thames, & Phelps, 2008; Hill, Ball, & Schilling, 2008) in designing my interview tasks to elicit number theory PCK and again to analyze responses.

Methodology

This interpretive case study (Merriam, 1998) centered on preservice elementary teachers who were seeking a mathematics concentration and enrolled in a number theory course. The majority of class time in this course was spent working on problem sets collaboratively (which were occasionally geared towards elementary school applications of number theory), and the instructor encouraged basic explanations or picture proofs. Data for this study came from multiple sources: classroom observational notes, student coursework, as well as responses from two sets of one-on-one task-based interviews, which served as the focus of the data analysis. Constant-comparative coding (Corbin & Strauss, 2008) was used as part of the coding process. Among my efforts to ensure trustworthiness, I used member checking during the interviews and data triangulation afterwards.

The six interview participants' had varying coursework experiences prior to the number theory course. Brit had taken a number and operations course designed for non-mathematics concentration elementary education majors, which encouraged collaborative, non-directive, and discovery learning, oftentimes with pictures and manipulatives. Cara, Gwen, and Lucy had completed a mathematics education course for mathematics concentration students, with a similar emphasis. As part of this course, Cara, Gwen, and Lucy served as teacher aides in the number and operations course. For Eden and Isla, the number theory course was their first experience with number and operations since grade school. Participants had varying amounts of tutoring experience; Brit and Lucy had the most experience.

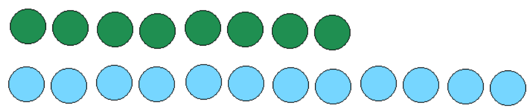
Results

Some interview tasks posed student scenarios to elicit participants' number theory PCK, so it was unsurprising that analysis of these tasks revealed themes of KCS, KCT, and knowledge of curriculum. It was also unsurprising to find that participants used SCK when responding to the content of these tasks. The degree to which participants' PCK *depended* on their SCK, however, was intriguing. Even more intriguing was the finding that this relationship between participants' SCK and PCK appeared to be stronger depending on participants' perspectives on how elementary school students learn mathematics. I discuss these emergent themes here by selecting a few interview tasks that elicited rich responses.

The tasks I discuss here all exemplify student-teacher interactions where a student presents an idea or conjecture and the teacher responds. Mathematics teachers spend a great deal of their time evaluating student work. This can range from checking answers to validating new conjectures and alternative algorithms, all of which classifies as SCK. As an example, consider the interview task in Figure 1. In this scenario, Maria is using colored chips to determine the GCF of 8 and 12, only to discover that the difference of 8 and 12 is also their GCF. Participants were asked to validate Maria's conjecture that the difference of two numbers is the same as their GCF.

Figure 1. *Interview Task 1: Student conjecture for finding GCF*

You have given each of your students 8 green chips and 12 blue chips and asked them to use the chips to find the GCF of 8 and 12. When she paired up the green and blue chips, Maria, noticed that there were 4 blue chips left.



As a result of this observation, Maria then tried making groups of 4 green chips and 4 blue chips. When she was successful, Maria conjectured that the difference between any two numbers is also their GCF.

(a) Is Maria's conjecture valid? How do you know?
(b) Is the GCF of two numbers related to their difference at all? How do you know?

In response to part (a), all six participants determined that this conjecture did not always work, because they were each able to find a counterexample. Having found a counterexample, a teacher may merely respond to the student that the conjecture is invalid. However, to give credence to Maria's conjecture and build on her innovation, a teacher may also recognize that the GCF divides the difference, i.e., that the difference can be broken into equal groups the size of the GCF, which can help Maria to find the GCF and simultaneously build on her conjecture. Thus, participants were also asked to explore Maria's conjecture further in part (b).

Although the participants had explored and used the idea that the GCF of two numbers divides their difference on numerous occasions in their number theory course, only one participant precariously recognized the relationship after producing and comparing several counterexamples. Other participants, like Isla, went as far to say that since there was no clear relationship between the difference of two numbers and their GCF, she would discourage Maria from thinking this way because "it could backfire… and you'd just get into trouble." While the task was not meant to elicit KCT, most participants suggested that they would convince Maria that her conjecture was incorrect and discourage her from using it. A stronger response might be to encourage Maria to investigate the concept further and guide her to a deeper realization about the content. It was unclear here whether participants' own understanding of the relationship between numbers' GCF and difference hindered their KCT, but it was more evident in other tasks.

While the previous interview task illustrates an obscure student conjecture, teachers still need content knowledge to evaluate and respond to more common student misconceptions. During another interview task, I asked participants to validate and respond to Mark, another hypothetical student, about his conjecture that the product of two numbers is also their LCM. All participants determined that the conjecture was incorrect by generating a counterexample, but they also determined that Mark's conjecture was reasonable since it occasionally worked. Two participants suggested that Mark's conjecture may have resulted from the observation that products of small numbers (with which elementary school students are most familiar) are quite often their LCMs. Lucy went further to suggest that Mark's familiarity with area models could have contributed to this misconception. She reasoned that since an A by B rectangular array can be broken into A

groups of B objects and B groups of A objects, showing that the area, or product, is a multiple of both A and B. Participants' attempts to understand why Mark might believe his conjecture exhibited KCS. Three of the participants' responses drew on knowledge of curriculum.

To elicit KCT, I also asked participants how they might respond to Mark to help him understand his misconceptions. To respond to Mark in a way that best built on his understanding, participants need to understand when and why Mark's conjecture occasionally works. Participants had explored the special properties of relatively prime numbers and the relationships between numbers' LCMs and GCFs at length in their number theory class, but many of them struggled to connect this task to their course work. Only three participants stated with confidence that Mark's conjecture only worked with pairs of relatively prime numbers, two of whom explained using the relationship between the GCF and the LCM, and one of whom also provided a nearly accurate explanation using mods. The other three participants claimed that Mark's conjecture only worked for pairs of prime numbers (Gwen claimed that consecutive integers would also work). Two of these participants explained that these pairs "don't have anything in common", implying that they do not have any common factors.

In responding to Mark, participants drew from their evaluation of his conjecture, depending on their SCK. Four participants suggested that Mark investigate predetermined pairs of numbers that would lead him to realize the conjecture was false. Three of these participants also suggested that they would encourage him to use manipulatives, and two of the participants commented that they would emphasize that Mark's answer would be a common multiple if not the LCM. While none of these responses outwardly exhibits weak KCT, participants suggested that they would pick pairs of numbers based on the types of counterexamples and examples that they recognized, possibly leading Mark to make other conjecture about which types of numbers worked.

This may have been counteracted with encouraging Mark to further explore on his own, but Eden and Isla, the two participants that had not taken the mathematics education course or the number and operations course, said they would demonstrate a counterexample for Mark using manipulatives. Eden and Isla explained that many students are visual learners and would benefit from this demonstration, while the other participants emphasized the value of learners discovering their own misconceptions. Interestingly, even though Cara and Lucy demonstrated an understanding of the relationship between GCFs and LCMs, they did not encourage Mark to explore this connection or adapt his conjecture.

Even more prominent of a misconception in elementary school number theory (and beyond) than Mark's conjecture is the idea that 1 is a prime number, the focus of a third interview task. This is a reasonable conclusion when considering the common 'definition' of *prime*: a number is *prime* if it is only divisible by 1 and itself. Many elementary school curricula add that prime numbers are *greater than* 1, but all six participants neglected this caveat when asked to define *prime*. Although this misconception was discussed in participants' number theory class (participants explored reasons why 1 could not be prime and ways for distinguishing it from prime numbers), when asked about it during an interview task some participants' still waivered in their reasoning. Half admitted to thinking that 1 was prime prior to the number theory class, recalling that their elementary school teachers *taught them that 1 was prime.*

When asked why 1 could not be prime, half of the participants claimed that factor trees would never end if 1 were prime. Gwen, however, was not convinced by this argument because the branches of a factor tree were supposed to stop once you reached a prime number. She and Isla were more comfortable with thinking of prime numbers as having exactly two factors, which would mean that 1 was not prime by definition. Cara reasoned that 1 was a square number and that square numbers were not prime. Lucy reasoned that the prime factorization of 1 could have infinitely many forms, which was “not alright”, but she neglected to fully address the fundamental theorem of arithmetic. All of the participants said they would use the same reasoning that had convinced them that 1 was not prime to convince an elementary school student of the same idea, thus confounding their KCT with insufficient SCK. A few of the participants were also unsure how to make their explanations grade level appropriate, exhibiting weak knowledge of curriculum.

Discussion

Each of the interviews tasks described above was designed to elicit number theory knowledge for teaching. Using Ball and colleagues’ framework, we see that participants used SCK by validating innovative student thinking. Whenever participants explained why a student might have a misconception about a concept, they were demonstrating KCS. And while the clinical interview setting did not allow for teaching demonstrations, participants’ descriptions for how they might help students better understand the material (e.g., using counterexamples or manipulatives) suggests they possessed knowledge of content and teaching (KCT).

Encouraged (to varying degrees) by their constructivist-style learning experiences and by personal tutoring experiences, all of the participants expressed an interest in establishing how students thought about a concept (KCS) and building on that understanding (KCT). It was clear from their responses to the interview tasks that participants drew from their SCK to do this. The participants with more constructivist experiences seemed more eager to encourage guided exploration rather than demonstrate a predetermined example, which requires less SCK. All participants acknowledged that they lacked or needed an advanced understanding of the material to best respond to student reasoning. This sentiment is exemplified by a thought from Brit:

> “[It’s important that I] know why it works and the different ways, so that I know the higher math (that I don't think [students] need to know), but I have the confidence that it works and why it works, so that if [students] really struggle, I can explain pieces of that about how it works every time and why it works.”

In general, I found that participants developed various pedagogical techniques with which to respond to students, e.g., guided-discovery. However, they lacked much of the SCK with which to apply it, in spite of their more abstract coursework. To better use their content knowledge in responding to students, preservice elementary teachers may benefit from further experiences or instruction to help them draw from their abstract understanding of the content to develop their SCK.

Results

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Brown, A., Thomas, K, & Tolias, G. (2002). Conceptions of divisibility: Success and understanding. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 41-82). Westport, CT: Ablex.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Thousand Oaks, CA: Sage.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers’ mathematics knowledge for teaching. *Elementary School Journal, 105*, 11-30.

Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education*, Jossey-Bass, San Francisco, CA.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15(2)*, 4-14.

Cobb, P. & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*(3/4), 175-190.

Zazkis, R. (1998). Odds and ends of odds and ends: an inquiry into students' understanding of even and odd numbers. *Educational Studies in Mathematics, 36*, 73-89.

Zazkis, R. & Campbell, S. (1996). Divisibility and multiplicative structure of natural numbers: preservice teachers' understanding. *Journal for Research in Mathematics Education Research, 27*(5), 540-563.

Zazkis, R. & Liljedahl, P. (2004). Understanding primes: The role of representation. *Journal for Research in Mathematics Education, 35*(3), 164-186.

THE CONSTRUCTION OF COHOMOLOGY AS OBJECTIFIED ACTION

Anderson Norton
Virginia Tech

The purpose of this paper is to investigate a theory about the nature of mathematical development, in which mathematics is characterized as the objectification of action. Informed by existing research on how students construct new mathematical objects, we consider as an example the psychological construction of cohomology and related objects of algebraic topology. This example extends neo-Piagetian theories of mathematical development from elementary school to graduate-level mathematics, while integrating existing research on students' learning of abstract algebra. Results of the investigation affirm the objectification of action as a distinguishing feature of mathematics in general, while indicating the kinds of mental actions that undergird the objects of advanced mathematics.

Key Words: Abstract Algebra, APOS Theory, Constructivism, Reflective Abstraction, Reification

"Mathematics is the science of actions without objects, and for that, of objects we can define through action." Paul Valéry (1973, p. 811).

When fields' medalist William Thurston endeavored to address the plight of mathematics education in the United States, he shared the following personal anecdote:

> I remember as a child, in fifth grade, coming to the amazing (to me) realization that the answer to 134 divided by 29 is 134/29 (and so forth). What a tremendous labor-saving device! To me, '134 divided by 29' meant a certain tedious chore, while 134/29 was an object with no implicit work. I went excitedly to my father to explain my major discovery. He told me that of course this is so, a/b and a divided by b are just synonyms. To him it was just a small variation in notation. (Thurston, 1990, p. 5)

Thurston used the story to illustrate the challenge we face, as teachers, when we attempt to unpack the mathematical objects we have constructed. Mathematics education researchers have taken pains to unpack the object of Thurston's example in particular, demonstrating how students begin to understand fractions (and especially improper fractions, like 134/29) as "numbers in their own right" (Hackenberg, 2007). The key to this and similar work has been to identify the mental actions that comprise those objects, thus equipping teachers and researchers with models for how students might construct those objects through activity.

Few students in the United States accomplish what Bill Thurston did (Norton & Wilkins, 2012). In fact, it's possible that Thurston's father did not appreciate his son's revelation because, for him, the fraction 134/29 symbolized nothing more than the division of two whole numbers. On the other hand, if the elder Thurston *had* constructed 134/29 as a number, it's probable that he would have forgotten the labor of that construction, which involves coordinating mental actions of partitioning and iterating within a three-level structure: 134/29 as a unit resulting from 134 iterations of a 1/29 unit, which results from partitioning a whole unit into 29 parts (Hackenberg, 2007). Figure 1 illustrates such a structure for the simpler fraction, 8/3. This structure supports a conception of the improper fraction as an object defined through its size relation with the whole: 8/3 as a number that is eight times as big at 1/3, which has a 1-to-3 size relation with the whole.

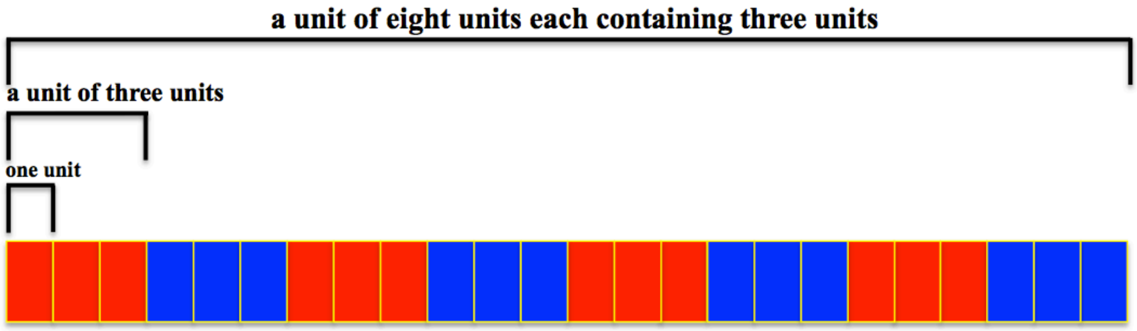


Figure 1. 8/3 as a unit of units of units.

Steffe and Olive (2010) have described this way of conceptualizing improper fractions as an *iterative fraction scheme* (IFS). Whereas we have fine-grained models for describing, explaining, and predicting the construction of improper fractions, few models of this kind exist for advanced mathematics. The scarcity of such models likely owes to two factors: (1) mapping the psychological construction of mathematics requires intensive and longitudinal studies of students' development—studies that, so far, have followed a trajectory from infancy to middle school mathematics; and (2) although schemes seem adequate for building models of development up to that point, modeling students' constructions of advanced mathematics likely requires more complex structures. Here, we will examine construction in an extreme case—cohomology—to identify key mental actions, even if we cannot model the complexity of their coordination.

Theoretical Framework

Inherent in Piaget's genetic epistemology is the idea that mathematical objects arise through the coordination of actions: "The meaning of objects has two aspects: It is 'what can be done with them' either physically or mentally… The meaning of object is also 'what it is made of,' or how it is composed. Here again, objects are subordinate to actions." (Piaget & Garcia, 1986, pp. 65-66). As Tall and colleagues (2000) have noted, several theoretical frameworks for teaching and learning have arisen from this idea, including APOS theory (Dubinsky, 1991), reification (Sfard, 1991), and scheme theory (von Glasersfeld, 1995). Here, we present a broader theoretical framework that builds on such work while aligning more closely with Piaget's characterizations of actions and objects, as well as his characterization of mathematics itself.

APOS Theory

Dubinsky and colleagues (e.g., Dubinsky & Lewin, 1986) developed APOS theory as a means of applying Piaget's constructivist epistemology to research on undergraduate mathematics education. In particular, they demonstrate how mathematical actions may become *reflectively abstracted* as advanced mathematical objects and schemas. Their central tenet is that "mathematical knowledge consists in an individual's tendency to deal with perceived mathematical problem situations by constructing mental *actions*, *processes*, and *objects* and organizing them into *schemas* to make sense of the situations and solve the problems" (Dubinsky & McDonald, 2001, p. 2). In this framework, *actions* are defined as transformations of tangible objects (including diagrams and written symbols) and might include carrying out the steps of an algorithm, such as computing the left cosets of a particular algebraic group. Reflecting on such actions allows the individual to internalize them as mental *processes* that the individual can imagine performing, without the need for tangible objects. Similar to Piaget (1970b), Dubinsky and McDonald (2001) argue that this internalization allows students to reverse and compose actions. The process becomes an *object* for an individual when he or she can symbolize it and purposefully act upon it. "Finally, a schema for a particular mathematical concept is an individual's collection of actions, processes, objects, and other schemas which are linked by some general principles to form a framework in the individual's mind" (p. 3).

Reification

Following Dubinksy (1986), Sfard (1992) further elaborated on Piaget's (1970a) notion of reflective abstraction by prescribing three stages through which students progress from engaging in mathematical processes to producing mathematical objects. To illustrate, Sfard provided an extended example from the historical development of number: from natural numbers, to positive rational numbers, to positive real numbers, to real numbers, and finally to complex numbers. She argues that each step-wise development has depended upon stages of *interiorization*, *condensation*, and *reification*. In particular, in the production of rational numbers, processes involving the division of natural numbers become *interiorized* so that they "can be carried out in mental representation" (p. 18, from Piaget, 1970a). Then they are *condensed* so that they can be combined with other processes, such as measurement. Finally, they are *reified*, or objectified, as a static structure on which to perform further processes, as in the development of positive real numbers. In fact, we can find evidence of this kind of development in the personal experience shared by Thurston: Whereas 129/34 had been a laborious process to perform, perhaps interiorized and condensed over a period of learning, in an instant it became reified as an object or "compact whole" (Sfard, 1992, p. 14). Unfortunately, Bill Thurston's father did not appreciate this "quantum leap" (p. 20) from process to object, which we might explain in either of two ways, as discussed later in this section.

Scheme Theory

Sfard did not make use of Dubinsky's *action-process* distinction, allowing processes to include actions, whether carried out physically or mentally. Neither did she make use of *schemas*. In contrast, scheme theory relies on a different characterization of action and utilizes a construct similar to Dubinsky's schema, but does not explicitly address the production of *objects*. von Glasersfeld (1995) described a *scheme* as a three-part structure: an assimilatory template of situations that might activate the scheme, a coordinated collection of mental actions carried out by the scheme, and an expected result from acting in the situation. Although Dubinsky's and Sfard's frameworks would include such actions, von Glasersfeld's description of mental action drew more heavily and narrowly from Piaget. For example, in contrast to the more formal mathematical actions of dividing and measuring described in Sfard's analysis of how students construct positive rational numbers, a scheme theoretic perspective would focus on the *psychological* actions that undergird them.

Actions and Objects

In an attempt to characterize the nature of mathematical objects and their construction, Tall and colleagues (2000) reviewed each of the frameworks described here and, noting the common theme of encapsulated actions, sought to describe how actions become objectified. Here, we broaden these frameworks and extend their purpose by arguing that mathematics *is* the objectification of action—this is what makes our field unique and, in some sense, infallible.

Unlike other sciences, languages, or any other field of study, all of the objects of mathematics are based on actions and their coordination so that, ultimately, mathematical claims are about nothing but the mental actions we can perform. If these actions correspond to (or even predict) experiential reality, it is only because we, as humans, have evolved to operate within the world we experience (Piaget, 1971/1970).

Piaget's epistemological research draws a fundamental distinction between two kinds of thought: figurative and operative. Whereas figurative thought pertains to empirical abstractions of "perception, imitation, and mental imagery" (1970a, p. 14), operative thought is the domain of mathematics. It pertains to reflective abstractions of one's coordinated activity in the construction of mental actions and structures. Unlike figurative objects (such as colors and drawings), operative objects remain dynamic on the basis of the actions that comprise them and the structures that organize them. Moreover, constructing such objects opens new possibilities for action, so that mathematics continually builds upon itself in alternating layers of actions and objects. Figure 2 illustrates the basic character of operative thought.

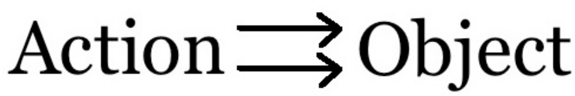


Figure 2. Mathematics as objectified action.

The top arrow in Figure 2 indicates that actions become reflectively abstracted as objects. The bottom arrow indicates that, as objects, these objectified actions can be acted upon. This pattern lies at the heart of Piaget's epistemology of mathematics and can also be found Sfard's reification and Dubinsky's APOS theory. What Sfard and Dubinsky do not address is how interiorized actions become organized within *psychological* (rather than formal mathematical) structures—the subject of Piaget's structuralism.

Structuralism

Structuralism focuses solely on operative thought, as an attempt to explain how children develop logico-mathematical reasoning. In addition to schemes (discussed above), Piaget (1970b) posited algebraic group-like structures that organize mental actions into reversible and composable systems. For example, students who have constructed mental actions of partitioning and iterating might organize them as inverse elements within a "splitting group", where iterating a part five times undoes the mental action of partitioning a continuous whole into five parts (Norton & Wilkins, 2012). They might also engage in recursive partitioning, in which partitioning is both an action and the object of that action (e.g., partitioning a continuous whole into three parts and then partitioning each of those parts into five parts to produce fifteen parts in the whole). Recent research (ibid) indicates that this group-like structure is necessary for the construction of IFS—the way of operating Thurston apparently constructed in fifth grade.

Although Piaget's epistemology (including his structuralism) equates logico-mathematical thought with operative thought, much of what happens in mathematics classroom involves figurative thought as well (Thompson, 1985). When the link is broken between a student's mental actions and the objects of a mathematical lesson, the student has little recourse but to engage in figurative thought. Sfard and Linchevski (1994) referred to this kind of engagement as the *pseudostructuralist approach*: "The new knowledge remains detached from its operational underpinnings and from previously developed systems of concepts" (p. 221). Moreover, Thompson (1985) has argued that students foreground some objects of mathematical discussion as operative—acting on them and deconstructing them into their constituent actions—while placing other objects in the background, as figurative. For example, functions might be operative in the context of high school algebra, as students act on covarying quantities and attempt to establish them as invariant relationships, but functions might be treated as figurative within cohomology, where they are elements of a group. In any case, what constitutes operative thought depends upon the available mental actions of the individual and her goals within the activity. Thus, we can say the same for mathematics.

Research on Abstract Algebra from an Action-Object Perspective

Action-object perspectives (especially Sfard's reification and Dubinsky's APOS theory) have gained strong influence in research on undergraduate mathematics education (RUME). Here, we review RUME studies from an action-object perspective that focus on concepts related to abstract algebra, and therefore related to algebraic topology and cohomology (for which no direct mathematics education research exists).

In a study on how college mathematics majors learn group isomorphism, Leron, Hazzan, and Zazkis (1995) drew a distinction between students who understood "the relation of two groups being isomorphic" and those who understood "the object of isomorphism" (p. 154). They identified three phases in students' transition from the former, action/process conception, to the latter, object conception: (1) concepts that reference the student doing something; (2) concepts that reference a process that could be carried out by anyone; (3) concepts that make claims of subject-independent existence. As students struggled to progress toward an object conception of isomorphism, the researchers noticed them "craving for canonical procedures and their fear of loose or uncertain procedures, indeed, procedures with any degree of freedom" (p. 171).

In a similar study with high school teachers, Dubinsky, Dautermann, Leron, and Zazkis (1994) focused on the interconnected layers of objects within group theory—group, subgroup, coset, normality, and quotient group—and their dependency on existing concepts of set and function. The teachers tended to begin by treating groups as sets on which to act and only later considered the role of a binary operator (function) in defining groups as objects. In line with Leron, Hazzan, and Zazkis (1995), the researchers noted the need for a concept of isomorphism in order to construct "group as an equivalence class of isomorphic pairs [of sets and functions]" (Dubinsky et al., 1994, p. 290). They also found that teachers construct subgroups in parallel with groups, as functions with a restricted domain. However, the teachers were generally not successful in constructing quotient groups, which the researchers attribute to difficulty in objectifying the process of forming cosets—a prerequisite construction for treating cosets as elements of a group. This difficulty was associated with teachers' tendency to conflate normality and commutativity.

Hazzan (1999) found that undergraduate students deal with the complexity of abstract algebra by "reducing the level of abstraction" (p. 71). Students do this in three distinct ways: (1) by basing arguments on more familiar mathematical entities (such as sets, rather than groups); (2) by dealing with single elements within a more complex collection (for example, working with a representative element within a quotient group, rather than the quotient group itself); and (3) by reducing objects to the actions that comprise them. Although the three methods are closely related, the third method aligns most directly with an action-object perspective. In line with the study by Leron, Hazzan, and Zazkis (1995), students can reduce the complexity of an entity by imagining actions they can perform to build it up. For example, one student dealt with quotient groups, G/H, by referencing the imagined activity of taking all elements of the normal subgroup, H, and choosing an element from the group G by which to multiply them on the right.

Other studies have demonstrated the efficacy of an action-object perspective as a pedagogical tool (e.g., Asiala, Dubinsky, Mathews, Morics, & Oktac, 1997; Brown, DeVries, Dubinsky, & Thomas, 1997). For example, Asiala and colleagues (1997) reported on the effectiveness of an abstract algebra course that explicitly attended to students' progressive constructions of actions, processes, objects, and schema. In particular, they described an action conception of coset as one in which students could work with simple and familiar groups/subgroups to build the coset. Students progress to process conceptions of coset when they can imagine computing the products just as the student in the example provided above, from Hazzan (1999). Students can then progress to object conceptions, in which they do not need to focus on the actions of building the coset and instead act on the coset itself. Finally, a coset schema is formed as a network of actions, processes, objects, and schemas, by relating cosets to concepts of groups, subgroups, normality, and quotient groups.

The Construction of Cohomology

When considering the complexities of an advanced mathematical idea, diagrams can provide some indication of their organization. Specifically, Figure 3 represents various components of cohomology and their relationships. However, for most of us, these components and relationships remain figurative rather than operative because they do not symbolize mental actions that we perform, nor objects that we act upon. The situation is completely analogous to that faced by middle school students as they begin engaging in algebraic manipulation without reference to underlying mental actions. For example, students commonly solve equations of the form *ax*=*b* by subtracting *a* from both sides of the equation. Correcting students' behavior in these instances is unproductive in terms of supporting algebraic reasoning. We need to address the source of the problem, that algebraic manipulations should become a proxy for underlying mental actions on previously constructed objects.

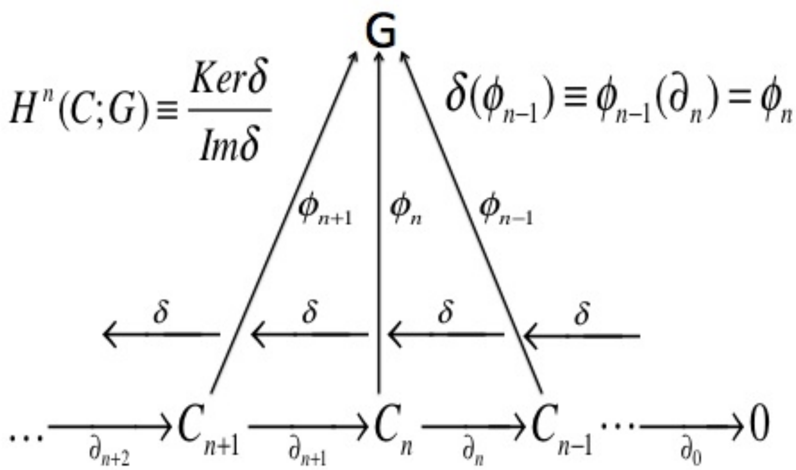


Figure 3. Diagram of cohomology

Previous research has suggested that constructing concepts in abstract algebra relies on having constructed functions and sets as objects first (Dubinsky, Dautermann, Leron, & Zazkis, 1994). Students tend to begin by treating groups as sets on which to act and only later consider the role of a binary operator (function) in defining groups as objects. Also, researchers have noted the interdependency of groups and isomorphisms in constructing “group as an equivalence class of isomorphic pairs [of sets and functions]” (Dubinsky et al., 1994, p. 290). Figure 3 begins at this stage, where C_n, represents a free abelian group generated by the set of n-dimensional triangles (e.g., vertices, edges, triangles, tetrahedras, etc.) used to build up the topological space under consideration. ∂_n represents a “boundary map” from C_n to C_{n-1}: a homomorphism that maps each n-dimensional triangle to its boundary (e.g., the boundary of an edge is the difference between its vertices, v_2-v_1). G represents another, selected group, and the various ϕs represent functions from C_n to G. Suppose these are objects for us, in the sense that Asiala and colleagues have described (1997): We can act on them and unpack them to their constituent actions (as opposed to figurative objects on which we might act but are not themselves composed of actions). Now consider the chain complex—the abelian groups, C_n, and the boundary maps, ∂_n, between them—as an algebraic procedure. Thus, Figure 3 serves to identify the boundary between algebraic objects and actions, even though we have not yet identified what psychological actions might undergird procedures associated the chain complex.

The Circle

To proceed, we might compute the homologies of familiar spaces. Computing homology allows us to focus on objectifying the chain complex while reducing further complexity introduced by *co*homology: the inclusion of the “ϕ” functions to group G and the coboundary maps, δ. Let us begin by computing the homology of the circle. This decision can be interpreted as an attempt to “reduce the level of abstraction” by dealing with a familiar entity (Hazzan, 1999), which might also make it easier to geometrically interpret the results of our algebraic computations. In particular, it is easy to see how a circle can be continuously deformed into a triangle, with three vertices and three edges. Thus, the chain complex becomes $0 \rightarrow \langle e_1, e_2, e_3 \rangle \rightarrow \langle v_1, v_2, v_3 \rangle \rightarrow 0$; that is, C_1 and C_2 are abelian groups generated by three elements and, thus, both are isomorphic to Z^3 (the product of three copies of the group of integers under addition). Now, the homology of the circle will be the quotient groups formed by the kernel of ∂_{n-1} mod the image of ∂_n.

Research indicates that constructing quotient groups is particularly challenging, even among students who have constructed groups as objects (Dubinsky et al., 1994). In the case of computing homologies, there is an additional challenge in making sense of the particular quotient groups defined by a particular homomorphism—the boundary map. Interpreting results geometrically gives these algebraic manipulations a geometric meaning, and the relevant mental actions lie therein. In other words, computing and interpreting homologies becomes a proxy for geometric actions associated with mapping n-dimensional triangles to their boundaries, equating sequences of n-dimensional triangles with an identity element, and forming n-dimensional loops around holes in the topological space under consideration. Thus, we begin to understand the chain complex as a representation of those actions. For the actions to become objectified, we need for them to define a class of spaces, so that homology becomes a proxy for that class.

In taking on this challenge, motivation quickly arises as a competing factor: Why did mathematicians ever bother to invent (co)homology in the first place? This as a competing factor because, for simple examples like the circle, sphere, or torus, there is no need for homology (let alone cohomology). We do not need to compute quotient groups of boundary mappings in order to determine that the torus and the sphere are topologically distinct. On the other hand, for the cases in which homology might be useful, the connection between the topology of the spaces and their homology (roughly, the connection between their geometry and their algebra) is opaque. We need to begin by working with simpler examples in order to build the connection in a way that might extend to ever more complex examples. Along the way, however, new complexities arise within the connection itself.

In working through examples, many of our actions will be conjectural—long sequences of tentative activity with depreciating confidence. For example, we might consider, "Why is homology invariant of choice of simplexes?" After all, we can build up the same topological space in many different ways. As it turns out, we do not even need to use n-dimensional triangles to form a chain complex, but can choose any n-dimensional polygon. Specifically, when computing the homology of the circle, we can choose any number, *m*, as the number of vertices (0-simplices) and edges (1-simplices). Figure 4 illustrates the cases of *m*=1 and *m*=3.

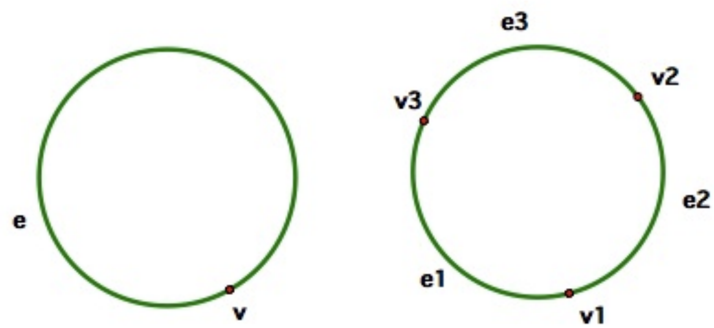


Figure 4. Two ways to form simplexes in the circle.

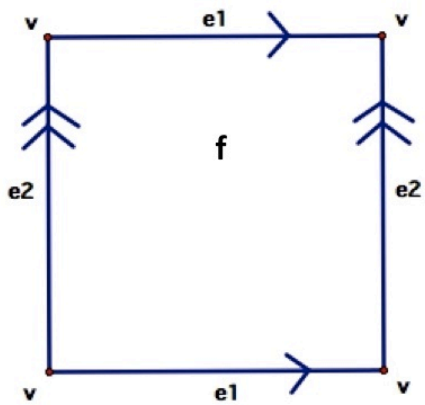
The image on the right of Figure 4 represents our original approach, with chain complex $0 \to Z^3 \to Z^3 \to 0$. The image on the left generates a simpler chain complex: $0 \to Z \to Z \to 0$. Even though the images and kernels within these mappings differ considerably, the resulting quotient groups are identical. For instance, in computing $H_0(X)$, the corresponding kernels are Z and Z^3, but the corresponding images are 0 and Z^2, so that the quotient group is Z in either case. Understanding why this happens is part of what it means to objectify the quotient groups that define homology. Just as understanding equivalent fractions involves more than showing that common factors cancel out, this understanding relies upon mental actions beyond the computation. Thus, the objectification of homology involves more than an interiorization of the boundary mapping or the process of computing its quotient groups. In particular, every time we add a new vertex to the simplex, we must add another edge, and the boundary of that edge will consist of two adjacent vertices. Their connectivity, as a single connected component, essentially leads to their identification in quotient group: Each vertex is identified with its two adjacent vertices, by the edge that connects them.

This understanding goes well beyond the process of computing kernels and images of the boundary map, and without this understanding, developed through simple examples, we would not be able to trust the extension of homology to the more complex examples where homology is actually useful. In building an understanding for how the algebraic computation of homology serves as a proxy from making topological distinctions, we find that relevant mental actions include geometric ones, related to vertex-edge graphs, as well as mental actions associated with continuity, especially as it relates to homotopy. By itself, the objectification of the boundary mapping would be no more useful to me than the algorithm for computing the product of two fractions; we would be objectifying something figurative rather than operative, and thus, would not be engaging in mathematics. We dig a little further into these actions by considering two nearly identical surfaces: the torus and the Klein bottle.

The Torus and the Klein Bottle

Topology is intended to address questions like the following: Are the torus and the Klein bottle continuous transformations of one another? Algebraic topology provides an answer by showing that the two surfaces have different homologies. Figure 5 demonstrates the homology of the torus.

$$0 \xrightarrow{\partial_3} < f > \xrightarrow{\partial_2} < e_1, e_2 > \xrightarrow{\partial_1} < v > \xrightarrow{\partial_0} 0$$



$$\partial_1(e) = v - v = 0$$

$$Ker(\partial_0) / \mathrm{Im}(\partial_1) = < v > /0 \cong \mathrm{Z}$$

$$\partial_2(f) = e_1 - e_2 - e_1 + e_2 = 0$$

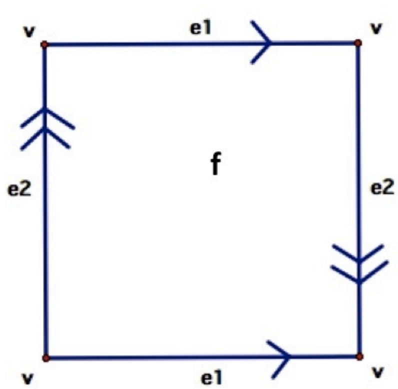
$$Ker(\partial_1) / \mathrm{Im}(\partial_2) = < e_1, e_2 > /0 \cong \mathrm{Z}^2$$

$$Ker(\partial_2) / \mathrm{Im}(\partial_3) = < f > /0 \cong \mathrm{Z}$$

Figure 5. Homology of the Torus

Note that the diagram on the left side of Figure 5 represents a torus because the opposite edges are identified with one another; i.e., we can produce the torus by gluing opposite edges together and, in the process, the four corners become a single vertex, *v*. Also note that each of the boundary maps turn out to be the 0 map because vertices and edges cancel out. Now consider the Klein bottle (Figure 6).

$$0 \xrightarrow{\partial_3} < f > \xrightarrow{\partial_2} < e_1, e_2 > \xrightarrow{\partial_1} < v > \xrightarrow{\partial_0} 0$$



$$Ker(\partial_0) / \mathrm{Im}(\partial_1) = < v > /0 = \mathrm{Z}$$

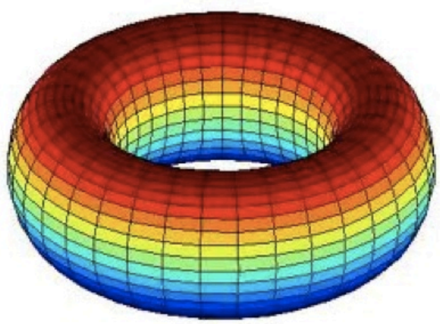
$$\partial_2(f) = e_1 + e_2 - e_1 + e_2 = 2e_2$$

$$Ker(\partial_1) / \mathrm{Im}(\partial_2) = < e_1, e_2 > /2 < e_2 > \cong \mathrm{Z} \times \mathrm{Z}_2$$

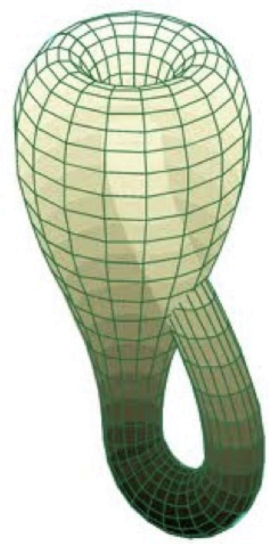
$$Ker(\partial_2) / \mathrm{Im}(\partial_3) = 0$$

Figure 6. Homology of the Klein Bottle.

The diagram (and therefore the homology) is exactly the same, except for one twist: A copy of e_2 is reversed. We can imagine both surfaces being constructed from a cylinder (after the pair of e_1s are identified), but in order to match up the directions of the two copies of e_2, the Kline bottle requires that the cylinder pass through itself to attach from the inside (see right side of Figure 7), which happens in four-dimensional space. Thus, the Klein bottle is a two-dimensional surface that does not exist in three-dimensional space. This fact alone might inform us that the torus and Klein bottle are not topologically equivalent, but we intend the comparison as an explanatory example for homology rather than a motivating one. We are trying to identify mental actions that might underlie our computations.



$$H_0 = \mathrm{Z}; H_1 = \mathrm{Z}^2; H_2 = \mathrm{Z}$$



$$H_0 = \mathrm{Z}; H_1 = \mathrm{Z} \times \mathrm{Z}_2; H_2 = 0$$

Figure 7. Homology as a proxy for topological actions.
Note: Images borrowed from Nosco (http://www.nosco.ch/mathematics/notes/torus.php) and Wikipedia (http://en.wikipedia.org/wiki/Klein_bottle) respectively.

In the case of the circle, we have already seen how the 0^{th} homology group, H_0, indicates the number of connected components in the topological space. Although the torus and Klein bottle affirm this connection (both are connected and have a single copy of Z for H_0), they do not provide interesting cases in this regard because we constructed each of them with only one vertex. However, they do provide an interesting contrast for H_1. How should we interpret the quotient groups Z^2 and ZxZ_2?

For both surfaces, the kernel of the 1^{st} boundary map (∂_1) is the group generated by the two edges; both of these edges form loops because their boundary is a single vertex, *v*, and for that same reason, they map to 0. For the torus, those loops are maintained when the face is glued on because the opposite edges match up. In order for them to match up, their directions must be opposite as we go around the boundary, and that is why they cancel out in the 2^{nd} boundary map (∂_2). In other words, the 2^{nd} boundary map is 0 precisely because the opposite edges of the face match up. Thus, the image is 0; no paths become identified with 0 in the quotient; and the 1^{st}

homology group (H_1) is the group generated by the two loops. We can see these loops on the torus in Figure 7: One goes around the "inner tube" and one goes around the hole at the center of the torus.

For the Klein bottle, one of the loops is transformed when the face is glued on because one pair of opposite edges does not match up. Instead of canceling out, the edge is doubled, and the 2nd boundary map has an image of $<2e_2>$. Thus, any even number of trips around the corresponding loop will be identified with 0. We can see the corresponding geometry in Figure 7: Tracing a loop around the "neck" of the bottle is just as it was for the "inner tube" of the torus, but tracing the other way yields a loop that undoes itself on the second pass because the trace moves to the other side of the surface (from inside out, or vice versa).

In general, the *kernel* of a boundary map is generated by n-dimensional *cycles*, and the *image* of the next boundary map is generated by the n-dimensional *boundaries* of n+1 dimensional polygons. In fact, algebraic topologists refer the kernels and images as "cycles" and "boundaries," respectively. In the quotient groups that define homology, the boundaries are identified with 0. Geometrically, we can understand this as gluing the cycles together (often in intricate ways). However, we can get lost in the computation of cycles, boundaries, and their quotients without ever considering the geometric actions to which they refer, much as middle school students do when they "complete the square" without ever considering the geometric square they are completing. Whether we are completing squares, connecting vertices, or gluing faces on to loops, the mathematics is in the geometric action for which the algebraic manipulation is a proxy. Once these actions are objectified, they can be symbolized in a way that conveys meaning. In particular, the symbols in Figure 3 become more than figurative material; they become proxies for objects, and actions on those objects.

Concluding Remarks

In reflecting on the actions and objects of cohomology, a key distinction arises—one that Piaget vigilantly maintained in his studies of young children but one that becomes easier to overlook when considering advanced mathematics: The bases for construction of formal mathematical objects are not necessarily formal processes. The diagram presented in Figure 3 might implicate computing kernels and images of boundary maps as primary actions to objectify, but subsequent investigation indicates a wide network of mostly geometric actions to coordinate. This finding supports the Piagetian notion that mathematics is a product of *psychological* action and not simply the enculturation of formal processes developed in the history of mathematics.

APOS theory (Dubinski, 1991) and reification (Sfard, 1992) have contributed greatly to mathematics education by extending Piaget's notion of reflective abstraction to advanced mathematics. However, researchers tend to use these frameworks as pedagogical tools for supporting student mastery of formal procedures, such as computing quotient groups (Asiala, Dubinsky, Matthews, Morics, & Oktac, 1997), especially when *actions* and *processes* refer to formal procedures. Although computations and procedures are integral to mathematical development, we must explicitly attend to the mental actions that give them meaning in order to support operative (and therefore mathematical) knowledge, rather than figurative knowledge. In fact, Sfard herself pointed to the "pitfall" of figurative knowledge when she warned of *pseudostructuralist* approaches to knowledge and learning (Sfard & Linchevski, 1994), which are indicated in students' aversion to "procedures with any degree of freedom" (Leron, Hazzan, & Zazkis, 1995). In contrast, a structuralist approach to mathematical knowledge and learning focuses on the construction and organization of reversible mental actions (Piaget, 1970b).

Scheme theory (von Glasersfeld, 1995) adopts a structuralist approach but has its own limitations in modeling the development of advanced mathematics; namely, the simplicity of a three-part structure may not accommodate the complexity of advanced mathematical concepts. Although we are able to identify some of the mental actions that undergird cohomology, we do not have models for their organization. This may explain why we often revert to figurative representations of knowledge (e.g., Figure 3) when investigating the development of advanced mathematics.

Our investigation of cohomology supports the argument that mathematics, at all levels, can be characterized as the objectification of action. This is the defining feature of mathematics, which distinguishes it from all other languages and sciences. Understanding mathematics in this way also evokes a degree of empathy as we provoke our students to construct new objects through action. In Bill Thurston's case, the father did not appreciate his son's accomplishment in constructing improper fractions as "numbers in their own right" (Hackenberg, 2007) because he could not unpack the coordinated actions of that construct. Likewise, models for teaching and learning advanced mathematics are limited by our models of the mental actions that comprise the objects of advanced mathematics.

References

Asiala, M., Dubinsky, E., Matthews, D. M., Morics, S. & Okatc, A. (1997). Development of students' understanding of cosets, normality, and quotient groups. *The Journal of Mathematical Behavior, 16*(3), 241-309.

Brown, A., DeVries, D. J., Dubinsky, E., & Thomas, K. (1997). Learning binary operations, groups, and subgroups. *The Journal of Mathematical Behavior*, *16*(3), 187-239.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95-123). Netherlands: Kluwer Academic Publishers.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics*, *27*(3), 267-305.

Dubinsky, E., & Lewin, P. (1986). Reflective abstraction and mathematics education: The genetic decomposition of induction and compactness. *Journal of Mathematical Behavior, 5*, 55-92.

Dubinsky, E., & McDonald, M. A. (2001). APOS: Constructivist theory of learning in undergraduate mathematics education research. *The Teaching and Learning of Mathematics at University Level, 7*(3), 275-282.

Hackenberg, A. J. (2007). Units coordination and the construction of improper fractions: A revision of the splitting hypothesis. *The Journal of Mathematical Behavior*, *26*(1), 27–47.

Hazzan, O. (1999). Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics, 40*, 71-90.

LaBoskey, V. K. (2004). The methodology of self-study and its theoretical underpinnings. In J. J. Loughran, M. L. Hamilton, V. K. LaBoskey, & T. Russell (Eds.), *International handbook of self-study of teaching and teacher education practices* (pp. 817–869). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Leron, U., Hazzan, O., & Zazkis, R. (1995*).* Learning group isomorphism: A crossroads of many concepts. *Educational Studies in Mathematics, 29, 153-174.*

Norton, A., & Wilkins, J. L. M. (2012). The splitting group. *Journal for Research in Mathematics Education*, *43*(5), 557-583.

Piaget, J. (1970). *Genetic epistemology* (E. Duckworth, Trans.). New York: Norton.

Piaget, J. (1970). *Structuralism* (C. Maschler, Trans.). New York: Basic Books.
Piaget, J. (1971). *Psychology and epistemology* (A. Rosin, Trans.). New York: Grossman Publishers (Originally published in 1970).
Piaget, J., & Garcia, R. (1991). *Toward a logic of meanings.* Hillsdale, NJ: Erlbaum.
Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on process and objects on different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1-36.
Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification – the case of function. In E. Dubinsky & G. Harel (Eds.), *The Concept of Function: Aspects of Epistemology and Pedagogy* (pp. 59-84). Washington, Mathematical Association of America.
Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification—the case of algebra. In *Learning mathematics* (pp. 87-124). Springer Netherlands.
Steffe, L. P., & Olive, J. (Eds.) (2010). *Children's fractional knowledge*. New York: Springer.
Tall, D., Thomas, M, Davis, G., Gray, E., & Simpson, A. (2000). What is the object of an encapsulation of a process? *Journal of Mathematical Behavior, 18*(2), 223-241.
Thompson, P. W. (1985). Experience, problem solving, and learning mathematics: Considerations in developing mathematics curricula. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 189–243). Hillsdale, NJ: Erlbaum.
Thurston, W. P. (1990). Mathematical education, *Notices of the American Mathematical Society, 37*, 844-850.
Valéry, P. (1974). *Cahiers, Vol. II*. Paris: Gallimard, 1974.
von Glasersfeld, E. (1995*). Radical constructivism: A way of knowing and learning*. London: Routledge Falmer.

BEYOND PLUG AND CHUG: THE NATURE OF CALCULUS HOMEWORK AT DOCTORAL INSTITUTIONS

Kady Hanson
San Diego State University

Gina Nunez
San Diego State University

Jessica Ellis
San Diego State University

Prior research reflects a positive relationship between homework and student academic achievement in undergraduate mathematics courses. Additionally, recent research has indicated no significant difference in student learning based upon the medium of the assignment (on-line based versus paper-based). These findings led us to ask the question: How does the nature of Calculus I homework assignments at doctoral institutions with successful calculus programs compare to assignments at institutions with less successful calculus programs? Descriptive analyses of student and instructor responses from a large national survey given to mainstream Calculus I programs were conducted. Analysis revealed significant differences in the nature of homework between successful and less successful institutions, including differences in the content and frequency of assignments. The holistic approach to homework taken by successful institutions adds to the existing literature on homework at the undergraduate level and indicates an interesting relationship between homework and student success in Calculus I courses.

Keywords: Doctoral Institutions, Calculus, Homework, Student Success, Quantitative Analyses

This study investigates the nature of homework assignments in Calculus I at doctoral institutions and their relationship to student success in Calculus I at these institutions. The effectiveness and delivery of assigning homework in order to promote student learning in mathematics at the undergraduate level has been investigated over many years (Cartledge & Sasser, 1981; Lenz, 2010). This research has revealed that homework assignments can have a positive effect on academic achievement when assigned and evaluated (Cartledge & Sasser, 1981). As innovations such as on-line homework systems emerge into the educational sector, the effect of homework on student learning continues to be investigated. For example, recent research found no significant difference in student learning whether homework was assigned on-line versus traditional paper-based homework with similar content (Lenz, 2010). Thus it appears that the medium of homework is less important than feedback or the content of assignments. In this report we address the following question: How does the nature of Calculus I homework assignments at doctoral institutions with successful calculus programs compare to the homework at those institutions identified as having less successful calculus programs?

Methods

The data for this study comes from a large national study of Calculus I programs. The study consisted of two phases, the first of which was a national survey given to calculus students and their instructors at the beginning and end of the term. The second phase of this study included case studies at five doctoral granting institutions deemed to have successful calculus programs as measured by increased student confidence, enjoyment, and interest in mathematics, Calculus I grade, and persistence onto Calculus II. This poster presentation will report on analyses of the end of term survey data from 3,187 students and 231 instructors as well as student focus group interviews at the five case study sites. Of the 3,187 students, 855 came from a case study institution – and thus a more successful institution. There were 231 instructors who completed the end of term survey, 49 of which came from a case study institution. Descriptive analyses were conducted on both student and instructor responses to understand the nature of the homework at successful and non-successful institutions.

Sample Results

As shown in Table 1, there were significant differences between student reports of the nature of homework assignments at successful versus less successful calculus programs. Compared to students in the less successful calculus programs, students at successful institutions report that the following happened more frequently: (a) assignments were assigned and collected, (b) the homework was returned with helpful feedback, and (c) students worked together on homework.

Table 1. Student reports of the nature of the assignments.

	Less Successful	Successful
1. How frequently did your instructor? *(1 = Not at all; 6 = Very often)*		
Require you to explain your thinking on your homework?***	3.08 (1.79)	3.53 (1.76)
2. Indicate how often the following occurred. *(1 = Never; 5 = Every class session)*		
My instructor assigned homework. **	4.06 (1.02)	4.33 (.95)
Homework was collected (either hard copy or online).*	3.41 (1.33)	3.82 (1.40)
3. Assignments completed outside of class time were: *(1 = Not at all; 6 = Very often)*		
Completed and graded online. ***	3.79 (2.29)	4.78 (1.92)
Graded and returned to me. **	4.36 (2.00)	4.17 (2.00)
Returned with helpful feedback/comments. ***	2.78 (1.79)	3.15 (1.84)
Submitted as a group project. ***	1.42 (1.09)	2.86 (2.02)
Challenging but doable.**	4.23 (1.36)	4.34 (1.27)
4. The assignments completed outside of class time required that I: *(1 = Not at all; 6 = Very often)*		
Solve word problems. ***	4.57 (1.25)	4.95 (1.13)
Solve problems unlike those done in class or in the book. ***	3.80 (1.63)	4.11 (1.53)
Use technology to understand ideas.***	2.50 (1.59)	3.14 (1.69)
5. Did you meet with other students to study or complete homework outside of class?* **		
Yes	57.6%	72.1%

Note.* = $p \leq .10$, ** = $p \leq .05$, *** = $p \leq .001$; Standard deviation in parentheses.

Analysis of survey data from instructors at doctoral institutions indicates a significant difference between the content of the homework assignments. Specifically, at institutions with successful calculus program a statistically significantly greater percentage of assigned problems focused on graphical interpretation, non-trivial or novel word problems, and proofs or justifications, as shown in Table 2. Ongoing analysis of focus group interviews at the five case study institutions reveals that students report assignments are mathematically challenging, faculty are supportive, and ample resources are available when assistance is needed.

Table 2. Instructors' End of Term Survey Response

On a typical assignment, what percentage of the problems focused on:	Less successful	Successful
Skills and methods for carrying out computations (e.g., methods of determining derivatives and antiderivatives)?***	51.15 (19.19)	39.80 (20.46)
Graphical interpretation of central ideas? ***	21.19 (11.79)	33.75 (20.80)
Solving standard word problems? *	23.11 (11.64)	27.08 (15.97)
Solving complex or unfamiliar word problems? ***	15.71 (11.57)	27.71 (23.36)
Proofs or justifications?**	9.47 (9.09)	14.13 (17.84)

Note.* = $p \leq .10$, ** = $p \leq .05$, *** = $p \leq .001$; Standard deviation in parentheses.

The results reveal interesting differences in the nature of homework assignments at institutions with more successful calculus programs compared to those institutions with less successful programs. In addition to the survey analyses, the poster presentation will provide an in depth analysis of the focus group interviews at the five case study sites.

References

Cartledge, C. M., & Sasser, J. E. (1981). The effect of homework assignments on the mathematics achievement of college students in freshman algebra. Retrieved from http://www.johnsasser.com/pdf/article01.pdf

Lenz, L. (2010). The Effect of a Web-Based Homework System on Student Outcomes in a First-Year Mathematics Course. *Journal of Computers in Mathematics and Science Teaching,* 29(3), 233-246. Chesapeake, VA: AACE.

Title: Differentiated student thinking while solving a distance vs. time graph problem

Eric A. Pandiscio
University of Maine

Preliminary Research Report

Abstract:

This study probes the thinking of students at different stages: a) secondary students taking calculus, b) college students taking calculus, and c) college students pursuing teacher certification taking a mathematics course other than calculus. The study asks: 1) what is the nature of student thinking when solving a graph problem, and 2) do students with different levels of mathematical experience solve a graph problem differently? A pilot investigation reveals many students estimate answers, even if they had studied calculus. For the current study, data will be collected during Fall, 2013. Oral interviews will be conducted with a subset of the participants and coded via Grounded Theory (Strauss & Corbin, 1990; Dick, 2005). This work follows physics education (McDermott, Rosenquist & van Zee, 1987; Thornton & Sokoloff, 1990; Kim & Kim, 2005), and mathematics education (Chiu, Kessel, Moschkovich & Munch-Nunez, 2001; Moschkovich, 1996) that describe difficulties students have with graph interpretation.

Keywords:
Graph comprehension
Problem solving
Student thinking

This study probes the levels and types of thinking demonstrated by students at different stages of their mathematical studies. The target audience includes three distinct groups of students: a) secondary students enrolled in calculus, b) college students enrolled in calculus, and c) college students in a teacher preparation program who are enrolled in a college-level mathematics course other than calculus. The rationale for including calculus is that the mathematical task posed to students may be solved using calculus, although calculus is not necessary.

Specifically, the study asks the following research questions:

1) what is the nature of student thinking when solving a graph-based problem?
2) do students with different levels of mathematical experience solve a graph-based problem different from each other?

A pilot administration of the written task to a small sample reveals that many students estimate answers. Even those students who have studied calculus tend not to use calculus to identify an exact solution. The pilot also shows that some students do not distinguish between a variable rate of change and a constant rate of change. Finally, the pilot led the researcher to ponder why it is that few students utilize some basic ideas regarding geometric properties of the circle to help solve one aspect of the task. For the full study, data will be collected during the Fall, 2013 semester. Although the pilot only analyzed written work, the full study will include oral interviews will be conducted with a subset of the participants. These will be chosen based the nature of the written responses, with the intention of identifying categorically different solutions for the oral interviews. The goal is to gain insight into student thinking, such that future work may center around curricular and instruction approaches to enhance the range of tools students bring to graphical problem solving. Responses will be coded via a modified Grounded Theory (Strauss & Corbin, 1990; Dick, 2005). The rationale is to establish themes and patterns of student thinking across different levels of mathematical experience. In particular, focus will be on progression of thinking patterns at the collegiate level, and also on the manner in which those students who plan to teach secondary mathematics approach problem solving.

The problem task was chosen from outside a traditional mathematics topic to foster the sort of work that is recommended the National Council of Teachers of Mathematics which states that mathematics experiences at all levels should include opportunities to learn about mathematics by working on problems arising in contexts outside mathematics" (NCTM, 2000, pp. 65-66). Further, mathematical topics must not be taught in isolation but in conjunction with problem solving and with applications in real-world contexts" (Reys, et al 2009, p.19). The Conference Board on the Mathematical Sciences describes the need for major reform in the teaching of college mathematics to prospective teachers (one of the target audiences of the study); in particular the observation that in the past "learning mathematics has meant only learning its procedures" (CBMS, 2012, p. 11) and goes on to suggest that doing mathematics in ways consistent with mathematical practice will require a new emphasis on understanding and problem solving. Further, one of the major recommendations is for students to "engage in reasoning, explaining, making sense of the mathematics (CBMS, 2012, p. 17). Finally, since the researcher is concerned about student understanding, following the recommendation of the influential *How People Learn* (National Research Council, 2000), the task in the investigation was chosen because of the critical nature of context for transfer of learning, and it is important for students to study ideas through multiple contexts. In this case, rate of change is being explored through a graphical representation of a physical phenomenon.

The work in this study follows the tradition of studies within both physics education research (e.g. McDermott, Rosenquist & van Zee, 1987; Thornton & Sokoloff, 1990; Merhar, Planinsic, & Cepic, 2009; Kim & Kim, 2005), and mathematics education research (e.g. Chiu, Kessel, Moschkovich & Munch-Nunez, 2001; Moschkovich, 1996) that pursue the difficulties students have with graph interpretation. Much of the previous work has identified students making mistakes that have been described as either a "picture as graph" difficulty or involving confusion between slope and height, early data from the pilot show different sorts of difficulties with the given problem. One goal of this study is to elucidate and formalize these sorts of challenges seen in student work. A major aspect of the interviews is to gather more comprehensive explanations by the participants regarding why they included, or not, certain features on their graph than was visible in the written responses. As a fuller picture emerges of student difficulties, more information will be available to guide additional research aimed at resolving those difficulties.

The Graph Problem

> Imagine that you are going to walk along the inside lane of a 400 meter track. You will start at the midpoint of one of the straightaways, and you will walk at a constant speed for two laps, ending at the place where you started.
>
> A simplified diagram of a track is displayed in Figure 1. Please note that the straightaways are each 100 meters long, and the arc on each end is a semicircle that is also 100 meters in length.
> 1. Please sketch a graph showing the distance you are from the starting point vs. time. In this context, the distance is considered the shortest straight line from your location to the starting point.
>
> 2. Write a brief explanation of why you constructed the graph to look as it does.
>
> 3. Identify the point on the track where you will farthest (again, in a straight line distance) from the starting point.
>
> 4. Describe how you determined this point.

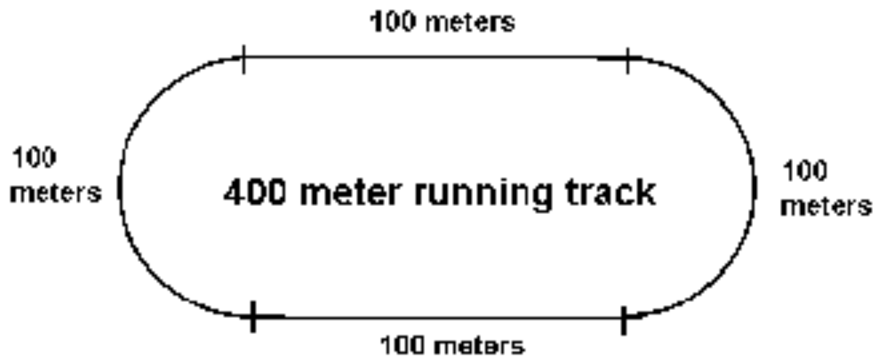


Figure 1

Preliminary questions for the audience:

1. How might I categorize student work differently such that I gain more insight into areas where different student difficulties overlap?

2. How do I extend findings from a single (hopefully robust) task towards suggestions for either: a) curriculum/instructional modifications, or b) a follow-up investigation with curriculum/instruction?

3. How likely is it that task-based interviews will reveal student reasoning in a graph-based problem solving context?

References

Chiu, M.M., Kessel, C., Moschkovich, J. & Munoz-Nunez, A. (2001). Learning to graph linear functions: A Case study of conceptual change. *Cognition and Instruction, 19*(2), 215-252.

Conference Board of Mathematical Sciences. 2012. The Mathematical Education of Teachers II. Providence, RI and Washington, DC. American Mathematical Society and Mathematical Association of America.

Dick, B. (2005). Grounded Theory: A thumbnail sketch. Retrieved from http://www.uq.net.au/action_research/arp/grounded.html

McDermott, L., Rosenquist, M. and van Zee, E. (1987). Student difficulties in connecting graphs and physics: examples from kinematics. *American Journal of Physics 55*(6), 502-513.

Merhar, V., Planinsic, G., and Cepic, M. (2009). Sketching graphs-an efficient way of probing students' conceptions. *European Journal of Physics, 30*, 163-175.

Moschkovich, J.N. (1996). Moving up and getting steeper: Negotiating shared descriptions of linear graphs. *Journal of the Learning Sciences, 5*(3), 239-277.

National Research Council (2000). *How People Learn*: *Brain, Mind, Experience, and School*. Bransford, J., Brown, A., and Cocking R., eds. National Academy Press. Washington, D.C.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, Virginia.

Reys, R., Lindquist, M., Lambdin, D., & Smith, N. (2009). *Helping Children Learn Mathematics* (Ninth Edition). John Wiley and Sons, Inc. Hoboken, NJ.

Strauss, A. & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park: CA: Sage.

Thornton, R., and Sokoloff, D. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics 58*, 858-867.
Van de Walle, J.A (2001). *Elementary and Middle School Mathematics: Teaching Developmentally, Fourth Edition*. New York: Addision Wesley Longman, Inc.

MATHEMATICS BEYOND CLASSROOM: STUDENTS' 'VALUE CREATION' THROUGH MATHEMATICAL MODELING WITHIN A LEARNING COMMUNITY

Joo young Park
Teachers College, Columbia University

This study examined how mathematical modeling activities within a collaborative group impact on students' perceived 'value' of mathematics. With a unified framework of Makiguchi's theory of 'value', mathematical disposition, and identity, the study identified the elements of the value-beauty, gains, and social good-with the observable evidences of mathematical disposition and identity. A total of 60 college students participated in 'Lifestyle' mathematical modeling project. Both qualitative and quantitative methods were used for data collection and analysis. The result from a paired-samples t-test showed the significant changes in students' mathematical disposition. The results from the analysis of students' written responses and interview data *described how the context of the modeling tasks and the collaborative group interplayed with students' perceived value. The poster will present the main findings and the examples of students' written tasks and responses.*

Key words: Mathematical Modeling, Instructional Activities and Practice, Value Creation

A number of studies demonstrated that mathematical modeling, which plays a prominent role in the new Common Core State Standards for Mathematics (CCSSM), promotes socially situated learning environments with group collaboration and creativity, and it has the potential to develops positive disposition toward mathematics and strengthen their mathematical identity (Ernest, 2002; Lesh & Doerr, 2003). This study involves inquires of what learning environment enables students to engage in meaningful mathematics learning and develop positive disposition as well as self-concept. The purpose of this study is to examine how mathematical modeling activities within a collaborative group impact on students' perceived 'value' of mathematics. The concept of 'value' was adopted from Makiguchi's theory of "value creation"(Bethel, 1989, p6).'Value creation' concerns with human development that enables individuals to gain benefits from developing a relationship with the object (mathematics) not only at the personal level but also societal level. With a unified framework of the theory of 'value', "mathematical disposition"(NCTM, 1989, p1), and identity, this study identified the elements of the value-beauty, gains, and social good-with the observable evidences of mathematical disposition and identity. A total of 60 students who enrolled in a college algebra course participated in 'LifeStyle' mathematical modeling project within a collaborative group. The topics of the modeling project were relevant to social and environmental issues in which students engaged in everyday lives. The result from paired samples t-test indicated the significant changes in students' mathematical disposition between pre and post survey. Based on the results from the analysis of students' journals and surveys, eighteen focal students were selected for interviews. The findings revealed that students develop an appreciation for mathematics as a useful and analytical tool to solve problems through engaging in the modeling project, and that participating in collaborative activities heightens students' interest and performance taking responsibility for mathematical meaning-making. Social value was created through students' interactions with the context of mathematical modeling and with peers while working in a group. The poster will present the main findings and the examples of students' written tasks and responses.

References

Bethel D. (1989). (Ed.) *Education for Creative Living: Ideas and Proposals of Tsunesaburo Makiguchi*, Iowa State University Press. An edit and translation into English of Makiguchi's work on pedagogy(1930).

Ernest, P. (2002). Empowerment in mathematics education. *Philosophy of Mathematics Education Journal* (15).

Lesh, R., & Doerr, H. (2003). Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp.3-34). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

111

STUDENT CALCULUS REASONING CONTEXTS

Matthew Petersen, Sarah Enoch, Jennifer Noll
Portland State University

This paper analyzes how student discourse about Calculus is situated in a graphical representation of a physics problem. Students were asked to identify three unlabeled graphs as representing the position, velocity and acceleration of a car. Findings showed that the students reasoned in three distinct contexts - static-graphical, covariational, and physical. While the students were able to communicate effectively between the first two contexts, and leverage them to find a solution to the problem, the students' discourse in the physical context did not communicate well with their discourse in the other two contexts, nor was it very fruitful in finding a solution to the problem.

Key words: [Calculus, Discourse, Physics, Reasoning Contexts]

Introduction & Background

Calculus has long been a gatekeeper for students entering STEM fields. Reform efforts geared to make calculus concepts more accessible to students have been underway now for multiple decades and so has research geared toward developing an understanding of student thinking about calculus. Much progress has been made but there is still much work to be done. This research study is part of a larger NSF-funded research study that continues systematic inquiry into both curriculum development and research on student thinking. This study investigates students from a newly developed calculus curriculum, Process-Oriented Guided Inquiry Learning (POGIL). The curriculum is activity based and student centered, and attempts to facilitate student construction of their own understanding.

The research to date reveals that foundational calculus concepts such as function, limit, derivative and integral are difficult for students and are often learned with significant misconceptions (Baker et al., 2000; Dreyfus & Eisenberg, 1981; Ferrini-Mundy & Graham, 1991; Tall & Vinner, 1981). Beyond identifying student conceptions and difficulties with calculus concepts, researchers have also investigated types of reasoning students use when thinking about calculus, and different theoretical perspectives (e.g., mental constructs, discourse) through which to view learning calculus. Researchers (Carlson et al., 2002; Zandieh & Knapp, 2006) investigating student thinking about derivative and related concepts using graphical tasks have identified different types of reasoning. For example, Carlson et al. identified the importance of covariational reasoning (the ability to coordinate the idea of a function's dependent variable changing with a given change in the independent variable). Zandieh and Knapp observed that students may reason about derivative from a number of contexts such as rate of change, velocity, slope of a tangent line, but that these various contexts may not be connected or deep for students. Ubez (2004) noted students have particular difficulty using the graph of a function to construct the graph of the derivative function.

Given that graphs and graphing play a key role in developing a deep understanding of calculus (and mathematics in general), it is important that students are able to process graphical representations of functions, derivatives and second derivatives, if they are to have a deep and well-connected understanding of calculus. Many of the studies cited above have made inroads into student thinking about graphs in relation to calculus ideas such as limit and derivative. Yet, more work is needed to better understand how student reasoning is situated within different contexts, how students make connections across different graphical contexts (e.g. covariational, or graphical—slope of a tangent line), as well as what sorts of thinking in these contexts supports or hinders learning. In this research study, we analyzed how students reasoned about a calculus task focused on a graphical representation of position, velocity, and acceleration of a moving car. The research questions guiding the work presented here are: (1) how do the students working in small groups on a calculus problem situated in the context of position, velocity, and acceleration communicate across contexts (static-graphical, covariational, physical, and equation), and (2) how do the various discursive contexts support or hinder their ability to reason about the task?

Theoretical Perspective

A commognitive framework (Sfard, 2008) is applied in the data analysis for the work presented in this paper. Sfard claims that mathematical discourses are distinguished by their objects, which arise in the discourse as what she calls a "realization tree", in which a signifier is potentially realized in a chain of realizations, each of which, from a different perspective, can

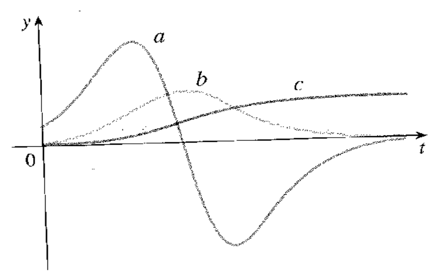


Figure 1. The figure above shows the graphs of three functions. One is the position function of a car, one is the velocity of the car, and one is its acceleration. Identify each curve (from Stewart, 2012).

also act as a signifier. These realizations can be signified by both verbal explanations, and physical gestures (e.g. Sfard, 2009). For the purposes of this study, we are considering the base signifier to be a problem-solving task which required the students to reason about position, velocity, and acceleration in the context of three unlabeled graphs (see Figure 1).

Sfard further argues that because the objects are discursive constructs, the realization trees are highly situated, and context specific. We have identified three contexts for student discourse: static-graphical, covariational, and physical.[1] Two of these three contexts, covariational and physical, have been identified in prior literature (Zandieh, 2000; Zandieh & Knapp, 2004; Carlson et al., 2002). The third context, static-graphical, is a new context developed through the current research. Each of these three contexts is discussed in the paragraph below.

For the first context, static-graphical, the realization tree consists of objects like derivative, slope, and concavity. Student discourse within this context is characterized by attention to features of the graphs, without treating the graphs as functions that are changing with respect to time. Such a discourse may be evident, for example, when a student places their hand horizontally along a graph to show where the slope is zero. The second context, covariational, is similar to the static-graphical context but the speaker attends to the graphs as if they were changing over time. In this second context a student may student move their hand along a graph

[1] A fourth context, equation, was also identified, but it is not addressed in this paper.

to show how the slopes of a graph are changing over time. To help distinguish between a static-graphical and covariational context, in a static-graphical context, students may characterize a curve as concave up, while in a covariational context they may characterize it as increasing but getting progressively steeper. In the physical context, students focus on aspects of the concrete physical motion, like velocity and acceleration, and students may use gestures such as moving their hand forward and backward to demonstrate physical motion.

It should be emphasized that we are not attempting to draw conclusions about students' internal mental schemes (e.g., Tall & Vinner, 1981), rather, we are attending to how a particular discursive context affects the students' ability to reason about a task. In this study, we analyze calculus students' ability to reason through a graphical problem-solving task. Our analysis focuses on the students' uses of the contexts described above, perceived as the realization trees for discourse around the root signifier (in this case, the given graphs and the corresponding problem). We investigate the students' abilities to communicate across these realization trees (i.e., across contexts), investigating to what extent their ability to communicate within and across these contexts influences their ability to reason about the task.

Methodology

Data for this paper were collected over the course of two terms from three sections of a ten week Calculus I course at a community college in the Pacific Northwest. The sections were all taught by the same instructor using the POGIL curriculum throughout the entire term. Students were recruited to participate in an interview during the ninth or tenth week of the term by flyers handed out during class and a follow-up email sent to those who had expressed interest. The volunteer students were interviewed in small groups of 3-4 students from the same class. During the interview, the students were given five Calculus problems to work on as a group. The students were asked to work together and to share their thinking out loud. While the interviewers occasionally probed the students to provide further explanation and gave prompts when it was appropriate for the students to move on to the next task, the interviewers, in general, tried not to ask many questions so that the students were allowed to work through the tasks fairly independently. The interviews were video-taped and the students were asked to record their work on a dry-erase board so that it could be visible on the video camera. The video recordings were subsequently transcribed. The results presented here are taken from their responses to the first of the five questions (see Figure 1).

Using the framework of the three discursive contexts, the data were analyzed identifying, for each line of reasoning a student offered to the group, which context(s) they were using to present their reasoning. Students' discourses were then analyzed with respect to whether or not there was interplay between the contexts and to what extent these discursive contexts were useful for reasoning through the task. At this time, some preliminary findings will be shared and discussed.

Results

Two groups addressed the question using a discourse based on a combination of the static-graphical and covariational realization of the functions. This was particularly striking for one group in particular. One student, Tim (a pseudonym), took the lead in analyzing the graph. During most of the interview, he focused the question on static-graphical or covariational features of the graph. His approach to answering the question is illustrated in the Table 1 below where he begins by explaining why graph **a** cannot be the position function:

Table 1. Student responses and corresponding gestures to the derivative task.

Speaker	Utterance	Gesture	Context
1. Tim	**a**'s not it [the position function], though.		
2. Lucas	It's not?		
3. Noelle	Why not?		
4. Tim:	'cause if **a**'s the position function, there's—derivative of this [graph **a]** is zero here [maximum of **a**] and zero here [minimum of **a**].	Places his hand horizontally along the maximum, and then minimum of graph **a**.	Static-Graphical
5. Lucas:	Which means that...		
6. Tim:	And there's no zeros on any of these curves [**b** or **c**].	Trances along the graph.	Static-Graphical
7. Tim:	But **c** could be it. See you can see when **c** is increasing [traces his finger along **c**] this [**b**] is positive as it starts to straighten out, **b** could be the derivative of that. And as you were talking [traces his finger along **c**] remember how you said there was kinda an inflection point here [on **c**]?	Traces his finger along **c** as he says that **c** is increasing, then along the tail end of **b**. Finally, runs his hand along **c**, in the neighborhood of its inflection point.	Covariational, transitioning into static-graphical when he mentions the inflection point.
8. Lucas	Right		
9. Tim	That's where the second derivative equals zero [points at the zero of **a**] that could be the inflection point and then this [**a**] is actually the second derivative.	Points to **a** as he first says "second derivative", then to its zero, then back at **a**.	Static-graphical

We coded utterances 4 and 9 in the transcript above as occurring in a static-graphical context because of Tim's focus on static features of the graphs: the extrema, inflection points, and zeroes of the functions. His horizontal gestures in utterance 4, and his pointing to specific locations in utterance 9, further support this conclusion. Conversely, we coded the beginning of utterance 7 as occurring in a covariational context because he attended to dynamic features of the graphs: "increasing" and "straightening out". At the end of utterance 7, he transitions back to a static-graphical context as he attends to the inflection point of graph **c**, a static feature of the graph, but his gesture tracing along the graph indicates that he may still be reasoning covariationally.

Another group took a very different approach to the question as their reasoning was heavily dependent on a physical context. The following episode in Table 2 is indicative of a physics-based approach to the question and the entire passage is identified as occurring in a physical context. The group had labeled graph **b** "function of car", and graph **c** acceleration, and Sally begins by arguing that the labels should be swapped:

Table 2. Student responses and gestures to the derivative task.

Speaker	Utterance	Gesture
1. Sally	I'm thinking that we should swap these two because the word in front of function car is position function of car. So it's moving, moving and then, you know, just stops moving, but it's moved.	Points to graphs **b** and **c** saying "these two". Then she gestures forward, palm down, as she says "moving moving.
2. Sally	And then acceleration, it's accelerating cause it's moving and then as it slows down and stops it would be decelerating. 5:42	Slowly moves her hand forward until "as it slows".
4. I	So wait [Sally] what are you proposing?	
3. Sally	Well, I'm thinking there's more to position function of the car; so it's not just the function of the car, it's position—as you know—it's moving forward. Well it's moving forward and it's not ever going backwards, and then accelerating,	She sharply moves her hand forward on "forward" and again on "moving forwards". On "backwards", she moves her hand back.
4. Sally	'Cause you know, start the car and then you're accelerating, accelerating, accelerating and you reach your top speed. And then you start slowing down, cause you're getting close to your destination or whatever. And then eventually you peter out and then you stop. But in the entire time you moved from point a to point b.	Traces her finger along graph **b**, beginning when she says "accelerating". Her finger is at the top of graph **b** when she says "top speed".

Sally's discourse around the task is taking place in a strictly physical context, as evidenced by her references to motion, acceleration, and speed, while not attending to any graphical features such as extrema, increasing or decreasing. This coding is further supported by her gestures which simulate the forward and backwards motion of a hypothetical car. In the interview from which this second example was drawn, the students frequently attempted to reason about the task within a physical discourse, as well as engaging in static-graphical and covariational discourses. However, they were not able to arrive at a reasonable solution within the physical discourse and eventually had to abandon this context in order to arrive at a reasonable solution.

Discussion and Conclusion

The examples given above illustrate two differing discourses (static graphical/ covariational and physical) that we observed as these groups of students attempted to reason through a problem-solving task involving graphical sense-making. The first group solved the problem by focusing on the static-graphical and covariational contexts for the problem, uniting the two into a single discourse, where each context informed and strengthened conclusions and ideas from the other. Translation from a static-graphical context to a covariational context seems to have occurred relatively fluently, and to have borne much fruit. The second group, on the other hand, attempted work in a physical context along with both covariational and static-graphical contexts. However, the use of various contexts seemed to have served more to confound their discourse, than to order and direct their communication. Perhaps because of this lack of constructive communication and also possibly because of the students' difficulty understanding acceleration, the use of a physical context bore relatively little fruit.

This research study is still in the preliminary stages of analysis. We will be interested in receiving feedback from our prospective audience and would like to hear their comments concerning the following questions: Do the three contexts described here adequately capture the situated discourse that takes place as students reason through problem-solving tasks around graphical representations of position, velocity, and acceleration? How can this analysis of students' reasoning within an interview be connected back to these students' experiences with the POGIL curriculum?

References

Baker, B., Cooley, L., and Trigueros, M. (2000). A Calculus Graphing Schema. *Journal for Research in Mathematics Education*, 31, 557-578.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., and Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 33, 352-378.

Dreyfus, T. and Eisenberg, T. (1982). Intuitive functional concepts: A baseline study on intuitions. *Journal for Research in Mathematics Education*, 13, 360-380.

Ferrini-Mundy, J., Graham, K. G. (1991). An overview of the calculus curriculum reform effort: Issues for learning, teaching, and curriculum development. *The American Mathematical Monthly*, 98, 627-635.

Sfard, A. (2008). *Thinking as Communicating: Human development, the growth of discourses, and mathematizing*. Cambridge, UK: Cambridge University Press.

Sfard, A. (2009). What's all the fuss about gestures? A commentary. *Educational Studies in Mathematics*, 70, 191-200.

Stewart, J. (2012). *Calculus, Seventh Edition*. Belmont, CA: Brooks/Cole, Cengage Learning.

Tall, D., Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, 12, 151-169.

Ubuz, B. (2007). Interpreting a graph and constructing its derivative graph: Stability and change in students' conceptions. *International Journal of Mathematical Education in Science and Technology*, 38, 609–637

Zandieh, M. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.) *Research in Collegiate Mathematics Education IV* (Vol. 8, pp. 103-126). Providence RI: American Mathematical Society.

Zandieh, M., Knapp, J. (2006). Exploring the role of metonymy in mathematical understanding and reasoning: The concept of derivative as an example. *Journal of Mathematical Behavior*, 25, 1–17.

112

STUDENT UNDERSTANDING OF LINEAR INDEPENDENCE OF FUNCTIONS

David Plaxco
Virginia Tech

Megan Wawro
Virginia Tech

Lizette Zietsman
Virginia Tech

In this study, we present preliminary findings regarding student understanding of linear independence of vector-valued functions. Students were given a series of homework questionnaires and participated in individual and paired interviews. The researchers used grounded theory to categorize student approaches for determining linear (in)dependence of functions. In order to gain insight into students' intuitive notions, data were collected before any formal instruction about the definition of linear independence of functions. The researchers describe initial analyses of student approaches, conjecturing their treatment of vector-valued functions at specific t-values or for varying t as a potentially beneficial lens of analysis. Students who evaluated specific t-values determined the linear independence of a set of vectors in $\mathbb{R}^2$ *rather than the linear independence of the set of functions, themselves elements of a function space. The analytical construct of process/object pairs (Sfard, 1991) could be a useful lens to explore this distinction.*

Key words: Linear Algebra, Differential Equations, Linear Independence, Function Space

Linear algebra and differential equations are important courses for mathematics and engineering students, and research shows that students tend to struggle with these courses (e.g., Dorier, 2000; Rasmussen, 2001). While research often focuses on student understanding of the mathematics in one of these courses, many topics (basis, Eigen theory, etc.) are integral to both. The focus of this study is how students make sense of linear independence of vector-valued functions, a concept common in linear algebra and differential equations. We focus on students' written responses regarding whether a given set of three functions from $\mathbb{R}$ to $\mathbb{R}^2$ is linearly independent, as well as two students' discussion of a similar question during a paired interview. These were intentionally given to students before formal instruction on linear independence of functions to reflect their initial notions of how linear independence might extend to functions. We describe initial analyses of student approaches, conjecturing their treatment of vector-valued functions at specific *t*-values or for varying *t* as a potentially beneficial lens of analysis. We close by addressing implications for teaching and by soliciting directions for future analysis and work.

Background and Literature

While a growing body of literature exists about student understanding of linear independence of vectors in $\mathbb{R}^n$ (e.g., Wawro & Plaxco, 2013; Bogomolny, 2007; Stewart & Thomas, 2010; Trigueros & Possani, 2011), we had difficulty finding empirical studies that report on student understanding of linearly independence of functions. In one report, however, Harel (2000) contended that a reason students have difficulty determining if the set $A = \{x, x^2, x^3, x^4\}$ is linearly independent is that the "students have not formed the concept of function as a mathematical object, as an entity in a vector space" (p. 181). Considering the definition of linear independence for vector-valued functions (see Figure 1), one can imagine that determining if there exists nonzero scalars that satisfy $a_1\overline{f_1}(t)$+ $a_2\overline{f_2}(t)$+…+ $a_n\overline{f_n}(t)$ = $\overline{0}$ for all *t* in *I* is a conceptual leap from determining if there exists nonzero scalars that satisfy $a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2$+…+ $a_n\boldsymbol{v}_n$ = $\mathbf{0}$ for "nonvarying" vectors $\boldsymbol{v}_i$ in $\mathbb{R}^n$. Indeed, our preliminary results are consistent with Harel's contention that student difficulty may rise from their struggle in treating the functions as objects in a vector space.

Instructors often show students the Wronskian when addressing ordinary differential equations because it provides an easy way to check if the solution functions are linearly independent on an interval. Although the Wronskian is a convenient tool, its use does not necessarily deepen students' understanding of linear independence of functions. Indeed, through personal communication, several differential equations instructors have expressed concern over students' insufficient understanding of linear independence of functions. This study strives to advance what is known in this area and inform subsequent pedagogical recommendations.

Let $\overline{f_i}: I = (a, b) \to \mathbb{R}^n$, $i = 1,2, \dots, n.$ The functions $\overline{f_1}, \overline{f_2}, \dots, \overline{f_n}$ are linearly independent on I if $a_i = 0$ $(i = 1,2, \dots, n)$ is the only solution to $a_1\overline{f_1}(t)+ a_2\overline{f_2}(t)+\ldots+ a_n\overline{f_n}(t) = \overline{0}$ for all $t \in I$.

Figure 1. Definition of linear independence of vector-valued functions.

Methods

Data were collected during the Fall 2012 semester in a course focused on core concepts in linear algebra and differential equations. Most students were first year, engineering or mathematics majors, and had scored a 4 or 5 on the Advanced Placement Calculus BC exam. Data sources included written work and video recordings of semi-structured individual and pairwise problem-solving interviews (Bernard, 1988). The written work was selected take-home questionnaires. The interviews were conducted at the midpoint and end of the semester. The key data sources for students' understanding of linear independence of vector-valued functions are the 4th and 5th questionnaires, the pairwise interview, and the second individual interview (see Figure 2). The questionnaires and pairwise interview occurred prior to students' encounter with the formal definition of linear independence of functions. We present preliminary analysis of Questions 2-3 (see Figure 2) in this proposal; analysis of Questions 1 and 4-7 is ongoing.

We used grounded theory (Strauss & Corbin, 1990) to iteratively analyze student responses to Question 2. We first described students' algebraic approaches and if they answered "linearly independent" (the correct answer), "linearly dependent," or something else. Descriptions were then iteratively compared for similarities and differences and grouped accordingly. This axial coding resulted in four categories of student approaches. Concurrently, we analyzed a pair's response to Question 3. We transcribed their interaction, watched the video, read the transcript, and drew quotes that helped support a base set of hypotheses about student thinking regarding linear independence of functions. The data from the paired interview (students discussing difficulties with each other, stating how they disagreed, etc.), which served as a counterpoint to the written data for Question 2, provided initial ideas of nuance regarding student thinking that the written data did not provide. This analysis informs our ongoing process of selectively coding (Strauss & Corbin, 1990) students' responses to Questions 1 and 4-7 (see Figure 2).

Preliminary Results

Twenty-four students' responses to Question 2 were sorted into four categories: Fix *t* First, Focus on Scalars, Function Combination, and Previous Rule (see Table 1). Initial analysis of the paired interview response to Question 3 provides insight into issues with linear independence of functions, namely, the difficulty in interpreting results of functions evaluated at specific *t* values.

Eight students' work indicated they approached Question 2 by evaluating the functions at specific values of *t* and comparing the resulting real-valued vectors; we categorize these approaches as "Fix *t* First." Six students using this approach concluded that the functions were linearly dependent because the image vectors were linearly dependent in $\mathbb{R}^2$, and two incorrectly determined that the image vectors were linearly independent in $\mathbb{R}^2$ (a correct response but with incorrect reasoning). Mathematically, all of the "Fix *t* first" solutions are incorrect because the students answered a different question than was posed. They drew conclusions about a set of vectors in $\mathbb{R}^2$ rather than about a set of functions in a function space.

Question from written homework questionnaire 4:

1. Consider the functions $F(t) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $G(t) = \begin{bmatrix} \sin t \\ 0 \end{bmatrix}$ with $F: \mathbb{R} \to \mathbb{R}^2$ and $G: \mathbb{R} \to \mathbb{R}^2$. Would you say these functions are linearly dependent for all $t \in \mathbb{R}$? Explain.

Question from written homework questionnaire 5:

2. "Consider the functions $F(t) = \begin{bmatrix} t \\ 1 \end{bmatrix}$, $G(t) = \begin{bmatrix} t^2 \\ 2 \end{bmatrix}$, and $H(t) = \begin{bmatrix} t^3 \\ 0 \end{bmatrix}$ with F: $\mathbb{R} \to \mathbb{R}^2$, G: $\mathbb{R} \to \mathbb{R}^2$, and H: $\mathbb{R} \to \mathbb{R}^2$. Would you say these functions are linearly dependent for all $t \in \mathbb{R}$? Explain your reasoning."

Question from pairwise interview:

3. "Consider the functions $F(t) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $G(t) = \begin{bmatrix} \sin t \\ 0 \end{bmatrix}$ with F: $\mathbb{R} \to \mathbb{R}^2$ and G: $\mathbb{R} \to \mathbb{R}^2$. Would you say these functions are linearly dependent or independent for all $t \in \mathbb{R}$? Explain."

Questions from end-of-semester individual interviews

4. Consider the functions $\overline{y_1}(t) = e^{2t} \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $\overline{y_2}(t) = e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Are these functions are linearly independent for all $t \in \mathbb{R}$? Explain.
5. Consider the functions $F(t) = \begin{bmatrix} \cos^2(t) \\ 1 \end{bmatrix}$, $G(t) = \begin{bmatrix} \sin^2(t) \\ 0 \end{bmatrix}$, and $H(t) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ with $F: \mathbb{R} \to \mathbb{R}^2$ and $G: \mathbb{R} \to \mathbb{R}^2$, and $H: \mathbb{R} \to \mathbb{R}^2$. Are these functions are linearly independent for all $t \in \mathbb{R}$? Explain.
6. Consider the functions $F(t) = \begin{bmatrix} t \\ 1 \end{bmatrix}$, $G(t) = \begin{bmatrix} t^2 \\ 2 \end{bmatrix}$, and $H(t) = \begin{bmatrix} t^3 \\ 0 \end{bmatrix}$ with $F: \mathbb{R} \to \mathbb{R}^2$ and $G: \mathbb{R} \to \mathbb{R}^2$, and $H: \mathbb{R} \to \mathbb{R}^2$. Are these functions are linearly independent for all $t \in \mathbb{R}$? Explain.
7. I think about linear independence of vector-valued functions the same way I think about linear independence of real-valued vectors.
   (Not at all) 1 2 3 4 5 (Very much)

Figure 2. Questions relevant to linear independence of vector-valued functions.

Seven students' work indicates that they focused primarily on the parameter scalars in the homogeneous vector equation (or corresponding system of equations), coded as "Focus on Scalars." Two of these students concluded that the set of functions was linearly independent, two concluded linearly dependent, two concluded linearly dependent for some values of *t* and linearly independent for others, and one could not draw a conclusion. While, overall, responses for this group were diverse, the common quality was that they focused on algebraic manipulation on the homogeneous equation with three scalar parameters (e.g., *a*, *b*, and *c*) and one variable, *t*. In most of these solutions, the student used the second components of the functions (or the second equation in the corresponding system) to eliminate one of the first two parameters and then re-wrote the first row (or first equation) to determine some relationship between the functions. None of these students evaluated the functions for a specific *t*-value but instead focused on the relationships that must hold true given a set of scalars that satisfy the homogeneous equation.

Six students provided solutions that are coded as "Function Combination." These students focused on whether t and t^2 could be combined to result in t^3, hence the name "Function Combination," and they all correctly concluded that the set of functions was linearly independent. Three of these students phrased their response as a relationship between vectors (written either verbally or algebraically), and the other three students phrased their response in terms of powers of *t*, that is, the first equation from the homogeneous system. Finally, three students' work was coded as a "Previous Rule" response. All three of these students stated that the set contained three vectors in $\mathbb{R}^2$ and, so, was linearly dependent. Note that given the set is a subset of the function space that contains functions from $\mathbb{R}$ to $\mathbb{R}^2$, not the vector space $\mathbb{R}^2$.

Table 1: Categories of Student Responses to the Question 2.

	Description	*Example*
Fix t First	Students evaluated the functions at one or more fixed value of *t*. Most then evaluated if the resulting real-valued vectors were LI or LD.	if $t=-2$ $\begin{bmatrix}-2\\1\end{bmatrix}$ $\begin{bmatrix}4\\2\end{bmatrix}$ $\begin{bmatrix}-8\\0\end{bmatrix}$ $2F(t)-G(t) = \begin{bmatrix}-4-4\\2-2\end{bmatrix} = \begin{bmatrix}-8\\0\end{bmatrix} = H(t)$
Focus on Scalars	Students carried out algebraic manipulations on the homogeneous vector equation (or related system of equations) without evaluating for specific t-values. They typically interpreted the results of their algebraic work by fixing the set of parameters and considering which/how many values of t satisfied the equation.	"It is LI at some values of t for a given a set of values a, b, and c but not at others" "No, because for any a_1, a_2, a_3, there exists a time t when the vectors are linearly independent, except for the zero vector." "There are plenty of t's where $aF + bG + cH \neq 0$ with the required values of a and b"
Function Combination	Students seem to attend to linear combinations of t, t^2, and t^3 of a variable *t* that is not constant. These students did not evaluate the functions for specific *t*-values.	"Can't change degree of t with scalar constants." "No linear combination of t and t^2 will ever yield t^3, so they are LI." "No because there is no way you can get one of the other vectors from a l.c. of the others. The only solution to $a_1F(t) + a_2G(t) + a_3H(t) = 0$ is $a_i = 0$, i = 1, 2, 3"
Previous Rule	Students relied on rules or generalizations about vectors in Euclidean spaces to determine linear (in)dependence of the set of functions.	"The set of equations would be L.D. because it has 3 components in $\mathbb{R}^2$." "Yes, *F*(*t*), *G*(*t*), and *H*(*t*) are LD b/c there are three vectors in $\mathbb{R}^2$."

In the paired interview, Jordan noted that the second components of the functions in Question 3 (Figure 2) required the scalar of the first function to be zero, but allowed the second scalar to be any real number. He then said that, because sin(*t*) is periodically zero, the second function was equal to the zero vector for some values of *t* – and so the vectors would be linearly dependent at those values of *t*. From this Jordan and Carter debated whether this meant that the functions were linearly dependent for all real values of *t*, or only at those specific values of *t*. Carter stated he needed a better understanding of what linear independence meant with respect to functions, pointing out that it might not be the case that "vectors being linearly dependent at some values" necessarily meant "linearly dependent at all values of *t*." Jordan, on the other hand, argued that linear dependence at any point in the interval meant that the vectors were linearly dependent on the entire interval. Later, when asked if they thought the functions were linearly dependent, the students responded simultaneously – Jordan, "Yes." and Carter, "No."

Discussion

In students' transition from linear independence of vectors in $\mathbb{R}^n$ to linear independence of functions in function spaces, we notice students' tendency to focus on specific t-values and consequently draw conclusions about vectors in $\mathbb{R}^2$ rather than about functions. The students whose work was coded as "Fix t First" drew conclusions about sets of vectors corresponding to specific t-values. Similarly, during the paired interview, Jordan focused on values of t for which the second function equaled the zero vector. While the sets of output vectors could be found to have a nontrivial solution to the homogeneous vector equation at specific t-values, the solutions to the homogeneous equation varied with t. All work coded "Function Combination" focused on the fact that no constant scalars could change the exponents of t, t^2, and t^3. To draw inferences about students' understanding of linear independence, we attend to the object on which a student acts. Students who evaluated specific t-values determined the linear independence of vectors in $\mathbb{R}^2$ while the "Function Combination" students likely considered the linear independence of the functions themselves. We hypothesize that the analytical construct of process/object pairs (e.g., Sfard, 1991; Zandieh, 2000) could be a useful lens to explore this distinction. For instance, realizing that "no constant scalars could change the exponents of t, t^2, and t^3" indicates at least a pseudo-object view of function. This highlights a need for students to begin to think of such functions as objects. Within a function space, linear independence may be defined (equivalent to the definition in Figure 1) according to a homogeneous equation in which the "zero" is the zero function rather than the zero vector. This indicates a need to modify the way that vector-valued functions are typically discussed during instruction so that functions may be viewed as an extension of the students' previous notions of linear independence to a new type of vector space.

In the talk, we will share updated analysis of Questions 1-7, teaching implications, and ask:

1. How could we explore hypotheses about student understanding of function and of formal logic accounting for difficulties with the definition of linear independence of functions?
2. Is the process-object lens fruitful / appropriate for categorizing student understanding of linear independence for real-valued vectors and for vector-valued functions?

References

Bernard R. H. (1988). *Research Methods in Cultural Anthropology*. London: Sage.

Bogomolny, M. (2007). Raising students' understanding: Linear algebra. In Woo, J. H., Lew, H. C., Park, K. S. & Seo, D. Y. (Eds.). Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education, Vol. 2 (pp. 65-72). Seoul: PME.

Dorier, J.-L. (Ed.). (2000). *On the teaching of linear algebra*. Dordrecht: Kluwer Academic Publisher.

Harel, G. (2000). Three principles of learning and teaching mathematics: Particular reference to linear algebra – Old and new observations. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 177-190). Dordrecht: Kluwer Academic Publisher.

Rasmussen, C. (2001). New directions in differential equations: A framework for interpreting students' understandings and difficulties. *Journal of Mathematical Behavior*, *20*, 55–87.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*, 1–36

Stewart, S. & Thomas, M. O. J. (2010). Student learning of basis, span and linear independence in linear algebra. *International Journal of Mathematical Education in Science and Technology*, *41*(2), 173-188.

Strauss, A., & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures*

and Techniques. Newsbury Park, California: Sage Publications.
Trigueros, M. & Possani, E. (2013). Using an economics model for teaching linear algebra. *Linear Algebra and its Applications, 438*, 1779-1792.
Wawro, M., & Plaxco, P. (2013). Utilizing types of mathematical activities to facilitate characterizing student understanding of span and linear independence. In (Eds.) S. Brown, G. Karakok, K. H. Roh, and M. Oehrtman, *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education, Volume I* (pp. 1-15), Denver, Colorado.
Zandieh, M. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. *Research in Collegiate Mathematics Education, IV* (Vol. 8, pp. 103-127).

113

LESSONS LEARNED FROM CASE STUDIES OF SUCCESSFUL CALCULUS PROGRAMS AT FIVE DOCTORAL DEGREE GRANTING INSTITUTIONS

[+]Chris Rasmussen, Jessica Ellis, [~]Dov Zazkis
[+]San Diego State University, [~]Rutgers University

In this report, we present initial findings from our case study analyses at five exemplary calculus programs at institutions that offer a doctoral degree in mathematics. Understanding the features that characterize exemplary calculus programs at doctoral degree granting institutions is particularly important because the vast majority of STEM graduates come from such institutions. Analysis of over 95 hours of interviews with faculty, administrators and students reveals seven different programmatic and structural features that are common across the five institutions. A community of practice and a social-academic integrations perspective are used to illuminate why and how these seven features contribute to successful calculus programs.

Keywords: Calculus, Student success, Case study

Calculus is typically the first mathematics course for science, technology, engineering, and mathematics (STEM) majors. Indeed, each fall approximately 300,000 college or university students, most of them in their first post-secondary year, take a course in differential calculus (Blair, Kirkman, & Maxwell, 2012). Nationally, there is a pressing need for students to be successful in calculus so that they can continue in their chosen STEM major and ultimately meet the growing demand of the workplace for STEM graduates (PCAST, 2012). However, student retention in STEM majors and the role of calculus in student persistence is a major problem (Rasmussen & Ellis, 2013; Seymour & Hewitt, 1997).

In order to better understand the terrain of the calculus teaching and learning in the United States, we are now in the fifth year of a five-year, large empirical study funded by the National Science Foundation and run under the auspices of the Mathematical Association of America. The goals of this project include: to improve our understanding of the demographics of students who enrol in calculus, to measure the impact of the various characteristics of calculus classes that are believed to influence student success, and to conduct explanatory case study analysis of exemplary programs in order to identify why and how these programs succeed. In this report, we present preliminary findings from our case study analyses at five exemplary calculus programs at institutions that offer a doctoral degree in mathematics. Understanding the features that characterize exemplary calculus programs at doctoral degree granting institutions is particularly important because these institutions produce the majority of STEM graduates.

The overall five-year project was conducted in two phases. In Phase 1 surveys were sent to a stratified random sample of students and their instructors at the beginning and the end of Calculus I. The surveys were restricted to “mainstream” calculus, meaning the calculus course designed to prepare students for the study of engineering or the mathematical or physical sciences. Surveys were designed to gain an overview of the various mainstream calculus programs nationwide, and to determine which institutions had more successful calculus programs. Success was defined by a combination of student variables: persistence in calculus as marked by stated intention to take Calculus II; affective changes, including enjoyment of math, confidence in mathematical ability, interest to continue studying math; and passing rates. In Phase 2 of the project, we conducted explanatory case studies at 18 different post secondary institutions, where the type of institution was determined by the highest degree offered in mathematics. In this report, we present findings from ongoing analyses of the five case studies at doctoral degree granting institutions.

Theoretical Background

Analysis of our case study data is grounded in two complementary perspectives, the first of which draws on the *community of practice* perspective put forth by Wenger and colleagues (Lave &Wenger; 1991; Wenger 1998). A community of practice is a collective construct in which the joint enterprise of achieving particular goals evolves and is sustained within the social connections of that particular group. In achieving a particular joint enterprise, such as the teaching and learning of calculus, a community of practice point of view highlights the role of brokers and boundary objects. A broker is someone who has membership status in more than one community and is in a position to infuse some element of one practice into another. The act of doing so is referred to as brokering (Wenger, 1998). Boundary objects are material things that allow people to cross between different communities and facilitate progress on their joint enterprise.

The second set of ideas that we employ to make sense of our case study data draws on research in Higher Education that has extensively studied factors related to student retention at the post-secondary level, with a focus on the effects of student engagement and integration on persistence (e.g., Kuh et al., 2008; Tinto, 1975, 2004). According to Tinto's integration framework (1975), persistence occurs when students are socially and academically integrated in the institution. This integration occurs through a negotiation between the students' incoming social and academic norms and the norms of the department and broader institution. From this perspective, student persistence (a measure of success in calculus) is viewed as a function of the dynamic relationship between the student and other actors within the institutional environment, including the classroom environment.

Method

The survey results from Phase 1 provided information on which institutions are enabling students to be more successful in Calculus I (as compared to other institutions of the same type) per our measures of success. Survey results, however well crafted and implemented, are limited in their ability to shed light on essential contextual aspects related to *why* and *how* institutions are producing students who are successful in calculus. The case studies were therefore designed to address this shortcoming by identifying and contextualizing the teaching practices, training practices, and institutional support practices that contribute to student success in Calculus I. As argued by Stake (1995) and Yin (2003), explanatory case studies are an appropriate methodology to study events (such as current practices in Calculus I) in situations in which the goal is to explain why or how, and for which there is little or no ability to control or manipulate relevant behaviors.

Four different case study teams (one per each type of institution—community college, bachelor, masters, and doctoral) conducted three-day site visits at the selected institutions. During the site visit each team, which consisted of 2-4 project team members, interviewed students, instructors, and administrators; observed classes; and collected exams, course materials, and homework. Common interview protocols for all 18 case studies were developed, piloted, and refined in order to facilitate comparison of calculus programs within and across institution type.

At the completion of each site visit the case study teams developed a reflective summary that captured much of what was learned about the calculus program, including key facts and features that were identified by both the case study team and the people interviewed as contributing to the success of the institution's calculus program. A more formal 3-4 page summary report was then developed by reviewing the reflective summary and transcripts and sent to the respective department of each institution as part of the member checking process (Stake, 1995).

At the five doctoral degree granting institutions, we conducted 92 interviews with instructors, administrators, and students for a total of more than 95 hours of audiorecordings. All interviews were fully transcribed and checked by a second person for accuracy and completeness. In order to manage this vast amount of qualitative data, a tagging scheme was developed to facilitate the location of relevant interview excerpts related to one of more of 30 different areas of interest. These areas of interest include such things as placement, technology, assignments and assessments, instructor characteristics, etc. Each interview was first chunked in terms of what we refer to as a "codeable unit." A codeable unit consists, more or less, of an interviewer question followed by a response. If a follow up question resulted in a new topic being discussed by the interviewee, then a new codeable unit is marked. Each codeable unit is then tagged with one or more of the 30+ codes. The idea is that once all interviews have been tagged with one or more codes, we can then systematically identify all instances in which any interviewee addressed a particular topic area. Once these instances have been located, then a more fine-grained grounded analysis will proceed. We used the facts and features documents to conduct initial cross case analysis. The grounded analysis will allow us to conduct cross institution case analysis arriving at an expanded list of important features.

The set of 30+ codes was developed by representatives from each of the four different case study teams and consists of both a priori codes from the literature and codes for themes that emerged from the reflective summaries. The final set of 30+ codes underwent an extensive cyclical process in which representatives from each case study team coded the same transcripts, vetted their respective coding, which then led to refining, deleting, and adding new codes and operational definitions. Two different team members coded each transcript and the two coders resolved any discrepancies.

Discussion

Cross case analysis of the five doctoral degree granting institutions led to the identification of seven features that contribute to the success of their calculus program. We first highlight *what* these seven features are followed by a discussion of the seven features in light of the communities of practice perspective and Tinto's academic and social integration perspective.

- **Coordination**. Calculus I (as well as PreCalculus and Calculus II) has a course Coordinator. The Coordinator holds regular meetings where calculus instructors talk about course pacing and coverage, develop midterm and final exams, discuss teaching and student difficulties, etc. Exams and finals are common and in some cases the homework assignments are coordinated.

- **Attending to Local Data**. There was someone in the department who routinely collected and analyzed data in order to inform and assess program changes. Departments did this work themselves and did not rely on the university to do so. Data collected and analyzed included pass rates, grade distributions, persistence, placement accuracy, and success in Calculus II.

- **Graduate Teaching Assistant (GTA) Training**. The more successful calculus program had substantive and well thought out GTA training programs. These ranged from a weeklong training prior to the semester together with follow up work during the semester to a semester course taken prior to teaching. The course included a significant amount of mentoring, practice teaching, and observing classes. GTA's were mentored in the use of active learning strategies in their recitation sections. The standard model of GTA's solving homework problems at the board was not the norm. The more successful calculus programs were moving toward more interactive and student centered recitation sections.

- **Active Learning**. Calculus instructors were encouraged to use and experiment with active learning strategies. In some cases the department Chair sent out regular emails with links to articles or other information about teaching. One institution even had biweekly teaching seminars led by the math faculty or invited experts. Particular instructional approaches were not prescribed or required for faculty at any of the institutions.

- **Rigorous Courses**. The more successful calculus programs tended to challenge students mathematically. They used textbooks and selected problems that required students to delve into concepts, work on modeling-type problems, or even proof-type problems. Techniques and skills were still highly valued. In some cases these were assessed separately and a satisfactory score on this assessment was a requirement for passing the course.

- **Learning Centers**. Students were provided with out of class resources. Almost every institution had a well-run and well-utilized tutoring center. In some cases this was a calculus only tutoring center and in other cases the tutoring center served linear algebra and differential equations. Tutoring labs had a director and tutors received training.

- **Placement**. Programs tended to have more than one way to determine student readiness for calculus. This included: placement exams (which were monitored to see if they were doing the job intended), gateway tests two weeks into the semester and different calculus format (e.g., more time) for students with lower algebra skills.

The fact that all five of the more successful calculus programs at doctoral degree granting institutions had someone whose official job included coordinating the different calculus sections is noteworthy. This role of coordinator was not something that rotated among faculty, such as committee assignments do, but rather was a designated and valued permanent position. The existence of this position is, however, only part of the story. An equally important part of the story is the role that calculus coordinator, among others, played in creating and sustaining a community of practice around the joint enterprise of teaching and learning of calculus. In the respective communities of practice, calculus was not seen as being under the purview of one person, such as the coordinator, but rather calculus was viewed as community property.

Nonetheless, the calculus coordinator played a unique role within their community of practice. In particular, the calculus coordinator functioned as a broker between the more central members in the department that typically teach calculus and the many newcomers. At doctoral institutions, these newcomers to the calculus joint enterprise include visiting research or teaching faculty, post docs, lecturers, and graduate teaching assistants (GTAs). The regular meetings that the calculus coordinator convened provided occasions for newcomers to be enculturated into the norms and practices related to calculus. We identified a number of boundary objects that helped to facilitate this enculturation, including historical records of passing rates, current grade and persistence data, student evaluations, various training manuals (especially for GTAs and visiting faculty) and the development of common assignments and assessments. Other brokers in the joint enterprise of teaching and learning calculus included, for some of the five doctoral institutions, the graduate teaching assistant trainers and leaders, department chair and the person whose responsibility it was to collect and disseminate to the department local data concern student pass rates and persistence and/or the correlation between these measures of success and the placement process. We conjecture that their attention to local data and continual improvement efforts contributed to a climate in which those involved with calculus teaching were always striving for improvement. Indeed, it was striking to us that none of the five case study institutions considered themselves to be particularly successful in calculus. That is, none of the five institutions in our case studies felt that they had everything just right.

A community of practice perspective helps to illuminate the how and why particular calculus programs are successful from a point of view that highlights faculty and administration. In our view, Tinto's academic and social integration perspective sheds equally important insight into how and why calculus programs are successful from a student point of view. In particular, almost without exception the students we talked with at the five doctoral institutions noted that they felt their calculus course was academically challenging (despite the fact that the vast majority had taken calculus in high school) but that there were a number of resources available to them to help them be successful. These resources included well-developed math help centers where students felt they received the help they needed and availability of instructor's and GTAs office hours. Other factors that contributed to students' academic and social integration included common space in the math department where students could gather to work on homework, dorms that provided them with opportunities to interact with like minded fellow students, and in some places a cohort system or strong student culture that provided cohesion between students.

In summary, our ongoing analysis of the five successful calculus programs at doctoral institutions is highlighting a number of structural and programmatic features that other institutions would likely to be interested in adapting. The ongoing theoretical analysis points to the importance of how these structural and programmatic features come together for faculty so that calculus is seen as community property and for the academic and social integration so critical for students' continued interest, enjoyment, and persistence in calculus. Our analysis that combines a community of practice perspective with the seminal work of Tinto on academic and social integration also sets the stage for the development of a more comprehensive model of successful college calculus programs.

References

Blair, R., Kirkman, E.E., & Maxwell, J.W. (2012), *Statistical abstract of undergraduate programs in the mathematical sciences in the United States*. *Conference Board of the Mathematical Sciences.* American Mathematical Society, Providence, RI.

Kuh, G., Cruce, T., Shoup, R., Kinzie, J., and Gonyea, R. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education,* 79 (5), 540-563.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
President's Council of Advisors on Science and Technology (PCAST) (2012). *Engage to excel: Producing one million additional college graduates with Degrees in Science, Technology, Engineering, and Mathematics*. Washington, DC: The White House.
Rasmussen, C., & Ellis, J. (2013). Students who switch out of calculus and the reasons why they leave. In Martinez, M. & Castro Superfine, A (Eds.). *Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 457-464). Chicago, IL: University of Illinois at Chicago.
Seymour, E. & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduate leave the sciences*. Boulder, CO: Westview Press.
Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.
Steen, L. A. (1988). *Calculus for a new century: A pump, not a filter. MAA Notes Number 8*. Washington, DC: MAA.
Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89-125.
Tinto, V. (2004). Linking learning and leaving. In J. M. Braxton (Ed.) *Reworking the student departure puzzle*. Nashville, TN: Vanderbilt University Press.
Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge, UK: Cambridge University Press.
Yin, R. (2003). *Case study research: Design and methods.* Thousand Oaks, CA: Sage Publications.

MEASUREMENT DEFINITIONS FOR ELEMENTARY SCHOOL TEACHERS: LINKS TO GRADUATE LEVEL MATHEMATICS

V. Rani Satyam
Michigan State University

Undergraduates planning to be teachers often encounter mathematics content textbooks written specifically for preservice teachers. Elementary mathematics textbooks of this kind provide in-depth definitions of elementary school mathematics to foster deeper understanding of these basic concepts. I looked at measurement definitions (length, area, and volume) across six preservice textbooks and identified overarching themes, using an open coding method. The following themes emerged across the set of definitions: discrete/continuous, unit, no overlaps/full cover, interior/exterior, function, measurement as an attribute, and space filling. This links to graduate level mathematics and has implications for preservice teachers and their future elementary students.

Key words: Measurement, Curriculum Analysis, Preservice Elementary Teachers

Definitions are central to mathematics, allowing for precise communication. Definitions also greatly determine how a person understands a concept (Tall & Vinner, 1981). Textbooks for teachers largely define the concepts and explanations of said concepts that a preservice teacher will pass on to their students (McCrory, 2006). Given this link, my research question is: What are the overarching themes in measurement definitions (length, area, and volume) in elementary mathematics textbooks for preservice teachers? The focus is on measurement, because studies show that young children struggle to understand fundamental measurement concepts (Battista, 2004; Outhred & Mitchelmore, 2000).

The data consisted of definitions of length, area, and volume across six elementary mathematics textbooks for preservice teachers. The textbook authors include professional mathematicians, and their background is reflected in the word choice of their definitions. Using grounded theory as a methodology, I employed an open coding strategy to identify frequently occurring keywords across the total set of definitions. The following characteristics emerged: discrete/continuous, unit, no overlaps/full cover, interior/exterior, function (associating one number to another), measurement as an attribute, and space filling. A discrete conception of measurement refers to units. For example, area may be defined as the number of units in a region. In contrast, a continuous conception refers to entire distance/space, no mention of units.

These themes are fundamental ideas of graduate level mathematics. The idea of interior/exterior links to analysis, the idea of a cover comes from topology, and the discrete/continuous difference is key in branches of applied mathematics. That these powerful ideas are at play in elementary school is promising, but unfortunately, they tend to be brought up informally or not mentioned at all. Preservice teachers may not explicitly think about these aspects in relation to measurement, because a concept like leaving no overlaps when measuring seems obvious. However, these aspects are not obvious to elementary students learning measurement for the first time.

The results of this curriculum analysis are an example of one way to analyze and characterize textbook definitions. The resulting themes also have implications for why teachers struggle to teach and students struggle to understand measurement, if there are radically different ways of explaining and thinking of measurement, depending on textbook.

References

Battista, M. T. (2004). Applying Cognition-Based Assessment to Elementary School Students' Development of Understanding of Area and Volume Measurement. *Mathematical Thinking and Learning*, *6*(2), 185–204.

McCrory, R. (2006). Mathematicians and Mathematics Textbooks for Prospective Elementary Teachers. *Notices of the AMS*, *53*(1), 20–29.

Outhred, L. N., & Mitchelmore, M. C. (2000). Young Children's Intuitive Understanding of Rectangular Area Measurement. *Journal for Research in Mathematics Education*, *31*(2), 144-167

Parker, T. H., & Baldridge, S. J. (2008). *Elementary geometry for teachers* (Volume 2). Okemos, MI: Sefton-Ash Publishing.

Tall, D., & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with Particular Reference to Limits and Continuity. *Educational Studies in Mathematics*, *12*(2), 151–169.

Mathematicians' views on transition-to-proof and advanced mathematics courses

Milos Savic, University of Oklahoma
Robert C. Moore, Andrews University
Melissa Mills, Oklahoma State University

This study explores mathematicians' views on 1) transition-to-proof courses, 2) knowledge and skills students need in order to succeed in subsequent mathematics courses, and 3) differences in the proving process across mathematical content areas. Seven mathematicians from three different universities (varying in department size), were interviewed. Precision, sense-making, flexibility, definition use, reading and validating proofs, and proof techniques are skills that the mathematicians stated were necessary to be successful in advanced mathematics courses. The participants agreed unanimously that a content course could be used as a transition-to-proof course under certain conditions. They also noted differences in the proving processes between abstract algebra and real analysis. Results from this study will be used to frame a larger study investigating students' proof processes in their subsequent mathematics content courses and investigating how these skills can be incorporated into a transition-to-proof course.

Key words: transition-to-proof, proof-based courses, mathematicians, proving process

Background Literature

Many mathematics departments in colleges and universities across the U.S. offer either a transition-to-proof course or a content course designated as a transition-to-proof course such as discrete mathematics or linear algebra. While there are mathematical topics that such courses commonly share, there is considerable variation in the mathematical content and the ways of teaching these courses. For example, some courses emphasize truth tables (in a limited sense) and logical reasoning explicitly (Epp, 2003), while others tend to focus on proving techniques using textbooks (e.g., Velleman, 1994). Some universities use a content course, such as linear algebra, abstract algebra, or real analysis as a transition-to-proof course. While students' proving and validation processes have been examined by researchers in different mathematical topics (Larsen & Zandieh, 2008; Alcock & Weber, 2005; Larson, Zandieh, & Rasmussen, 2008; Inglis, Mejia-Ramos, & Simpson, 2007; Lockwood & Strand, 2011), research studies on the content and effectiveness of those courses as a transition-to-proof course are few.

In one study, Alcock (2010) interviewed five mathematicians experienced in teaching a transition-to-proof course. She identified four modes of thinking (instantiation, creative thinking, critical thinking, and structural thinking) considered important by the mathematicians for successful proving and concluded that "it certainly seems reasonable to claim that collaborative classroom environments, in which students investigate, refine, and prove mathematical conjectures" (p. 94) foster the flexible use of all four modes.

Although Alcock and other researchers (e.g., Weber, 2010) discussed pedagogical strategies and implications and offered suggestions for teaching the four modes of thinking, they did not specify which mathematical topics can be useful in a transition-to-proof course for developing these four modes of thinking, nor did she address the question of whether transition-to-proof courses adequately prepare students for more advanced mathematics courses. The present study will begin to investigate these issues.

Research Questions

Specifically, the study seeks to extend the investigation of mathematicians' views regarding transition-to-proof courses by ascertaining their views on the knowledge and skills that students need in order to succeed in their *subsequent* advanced mathematics courses and how those knowledge and skills can be incorporated into a transition-to-proof course. Thus, the research questions for this study were:

1. What do mathematicians perceive as the necessary knowledge and skills that students need in order to succeed in their advanced mathematics courses?
2. What mathematical content do mathematicians consider to be appropriate for a transition-to-proof course?
3. What differences in proof and the proving process do mathematicians perceive across different mathematical content areas, and can these be incorporated into a transition course?

Methodology

Seven participants were interviewed from three different universities in the U.S. Professors A1 and A2 were from a small mathematics department (faculty of eight) located in the Midwest. Professors B1 (a mathematician researching in mathematics education), B2, and B3 were from a large mathematics department (faculty of 138) in the Midwest. Professors C1 and C2 were from a medium-sized mathematics department (faculty of 32) in the south central U.S. A comprehensive look at the backgrounds of all the professors is given in Table 1. The first row gives the professors' general area of study, the second row provides the professors' specific area of research, and the final row lists some of the courses the professors have taught in the last three years.

Table 1:

	A1	A2	B1	B2
General	Algebra	PDE	Math Education	Analysis, PDEs
Specific	Lie algebra	Nonlinear Elliptic	Development of Secondary Curriculum	Applications to Chemical Systems
Classes Previously Taught	Abstract, Linear algebra	Advanced calculus, Applied math	Transition, Math for elementary teachers, Geometry for teachers, Capstone course for secondary teachers	Calculus III, Foundations of applied math

	B3	C1	C2
General	Mathematical Physics	Topology	Analysis, Algebra, Math Ed
Specific	Random Schrodinger operations	Low-dimension	Finite groups
Classes Previously Taught	Transition, Honors real analysis, Capstone mathematics course	Abstract algebra, Geometry	Intro to real analysis, Intro to modern algebra, Linear algebra

The interviews included a series of questions pertaining to transition-to-proof courses and other advanced mathematics courses. The questions pertaining to this study that were asked in the interviews were:

1. What knowledge and skills do students need in order to be successful at writing proofs in advanced mathematics courses?
2. Do you think it is appropriate to teach a content course (e.g., linear algebra or analysis) as an introduction to proof course?
3. What are the differences between the proving process and proofs of certain mathematical topics, such as topology, algebra, and analysis?

The interviews were transcribed and analyzed by the authors using an open coding system.

Results

Knowledge and Skills

The mathematicians stated that there were a number of skills, listed in Table 2, needed in order to be successful at writing proofs. Each skill mentioned is accompanied by a representative quote from one of the mathematicians that claimed that the skill was necessary.

Table 2:

Knowledge and Skills	Participants	Representative Quote
Use of Definitions	A1,B1, B2, B3, C1, C2	C1: "I tell them over and over and over again, 'Definitions tell you how to write proofs.' …you look at the definition and that will tell you where to start."
Sense Making	A1, A2, B3, C1, C2	B3: "I've seen that light come on for some students when they do get it, when they realize what it's about. They make that transition from computational to proofing."
Learning Proving Techniques	A2, B1, B2, C2	B1: "What's an indirect proof? … What is proof by mathematical induction? How do you prove that two sets are equal? … So there are a variety of proof techniques that we introduce and talk about…"
Precision	A1,B2, B3, C1	A1: "So understanding what it means to be precise… rather than "I have a gut feeling and I think I understand how it works."
Reading/ Validating Proofs	A1, A2, C1	A1: "I think it's true that it should be a high priority to read a lot of proofs, and part of that is evaluating them."
Flexibility	A1, B2	B2: "*f(x)* if I do *f(y)* it's the same function…If I change the internal variable [*x* instead of *y*] [the students think that] everything is new."

Content course as a transition-to-proof course

All of the professors stated that teaching a content course (such as an upper-level undergraduate real analysis or abstract algebra course) could be possible as a transition-to-proof course only if the amount of content was drastically reduced and time was devoted to explicitly discussing proving techniques.

A1: "Yes, if the credits and time are increased enough to allow sufficient time to develop the habits of mind…I think [habits of mind do] need to [be] explicitly addressed, not just implicitly."

C1: "Yes. I absolutely believe so. But, I mean, you have to do it…so I can't teach a ring theory course [as a transition-to-proof course] with the idea that I have to cover a bunch of material."

Differences in proofs and proving between abstract algebra and real analysis

Mathematicians agreed that there were fundamental differences in proving between abstract algebra and real analysis and identified various aspects of those differences:

B1: "Well you have different definitions that are fundamental… In elementary analysis one needs to know, not only epsilon-delta definitions but all these tricks: given this epsilon we're gonna construct a delta using some sort of magic that comes from experience and you don't need that technique in an algebra class."

B3: "When I teach analysis I try to highlight certain things that are sort of very analysis-y. Like proving equality by proving two inequalities, which is not something you typically find in an algebra course."

C2: "I think the use of quantifiers is more difficult in analysis than in algebra… but I think that might possibly be offset a little by the fact that the analysis content area relates so solidly to the calculus that they have studied, so they have a good deal of computational experience."

Discussion

The seven mathematicians mentioned some of the knowledge and skills that have been discussed in the literature as problem areas in proving, such as the use of definitions to structure proofs, using mathematical language and notation, and having an intuitive understanding of the concepts (e.g., Moore, 1994; Edwards and Ward, 2004). It was interesting that the mathematicians, when asked about what skills students needed in proving in order to be successful, transitioned into discussing what proving skills their current or former students lacked.

According to the mathematicians, only a content course with enough time and reduced expectations on content could be used as a transition-to-proof course. However, the mathematicians voiced that there seem to be fundamental differences in the proving processes between abstract algebra and real analysis, including the use of quantifiers, difficult definitions, and the familiarity of the concepts. This dichotomy raises the question: How can a student transition to algebra proofs, for example, if he/she is exposed to proving when real analysis is used as an introductory proofs course? Finally, B1 and C2 mentioned quantifiers, which have been examined separately in the literature (Dubinsky & Yiparaki, 2000), while B3 mentioned proving equality by proving two inequalities. What other fundamental differences, not mentioned by the mathematicians, occur in separate courses? These fundamental differences can help with the development of curriculum for a transition-to-proof course that may improve the "transition" aspect.

Future Research

These results will be used to inform a larger study examining the proofs and proving process of students in their advanced mathematics courses. In the Fall of 2013 we will start a study examining the logic and proving process of students' proofs in different advanced mathematics content courses. Identifiable differences in proving across the content areas will inform the design of a transition-to-proof course that incorporates the proving techniques and the different modes of thinking from content courses that students will need to use in their subsequent content courses.

Questions

What would be other effective ways of surfacing fundamental differences between content? How can these fundamental differences be emphasized in content courses?

References

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Hoton, & P. Thompson (Eds.), *Research in Collegiate Mathematics Education, VII, CBMS Series Issues in Mathematics Education* (pp. 63-91). Providence, RI: American Mathematical Society.

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior 24*, 125-134.

Byrne, M. (2013). Cooperative learning and traversing the continuum of proof expertise: Preliminary results. *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education* (p. ??). Denver, CO: MAA.

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *CBMS Issues in Mathematics Education, Vol. 8: Research in Collegiate Mathematics Education IV* (pp. 239-289). Washington, D.C.: American Mathematical Society.

Edwards, B., & Ward, M. (2004). Surprises from mathematics education research: Student (mis)use of mathematical definitions. *MAA Monthly, 111*, 411-424.

Epp, S. S. (2003). The role of logic in teaching proof. *MAA Monthly, 110*, 886-899.

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modelling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66*, 3-21.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67*, 205-216.

Larson, C., Zandieh, M., & Rasmussen, C. (2008). *A trip through eigen-land: Where most roads lead to the direction associated with the largest eigenvalue.* Retrieved from Research in Undergraduate Mathematics Education Conference: ttp://sigmaa.maa.org/rume/crume2008/Proceedings/Larson%20SHORT.pdf

Larson, C., Zandieh, M., & Rasmussen, C. (2008). *A trip through eigen-land: Where most roads lead to the direction associated with the largest eigenvalue.* San Diego, CA: Paper presented at the Research in Undergraduate Mathematics Conference.

Lockwood, E., & Strand, S. (2011). Student use of set-oriented thinking in combinatorial problem solving. *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 132-135). Portland, OR: MAA.

Moore, R. (1994). Making the transition to formal proof. *Educational Studies in Mathematics 27*, 249-266.

Velleman, D. J. (1994). *How to prove it.* New York: Cambridge University Press.

Weber, K. (2010). The pedagogical practice of mathematicians with respect to proof. *Proceedings of the 13th Annual Conference on Research in Undergrdauate Mathematics Education.* Raleigh, NC: MAA.

A PROPOSAL FOR A THEORETICAL FRAMEWORK ON SPECIALIZED KNOWLEDGE FOR TEACHING MATHEMATICS

Thorsten Scheiner
University of Hamburg, Germany

Several approaches and models, partly distinct and partly overlapping, shape the theoretical landscape in mathematics education research on teacher knowledge. These approaches and models take Shulman's (1986) categorization of teacher knowledge, in particular 'subject matter knowledge' and 'pedagogical content knowledge', almost always as a point of departure. It seems safe to say that research on teacher knowledge has become more sophisticated in the sense that Shulman's dimensions of teacher knowledge are divided in sub-dimensions and looked at in more detail. For instance, subject matter knowledge can be further differentiated in terms of Schwab's substantive and syntactic structures, Harel's ways of understanding and ways of thinking, Bromme's school mathematical knowledge and academic content knowledge, or Ball and her colleagues' (Ball, Thames, & Phelps, 2008) common content knowledge and specialized content knowledge. Although each of these contributions shed light into important issues, among them it is only specialized content knowledge that can be considered as *unique* for the purposes of teaching mathematics. Building upon recent work addressing the 'uniqueness' of knowledge for teaching mathematics, the presented work provides a theory-driven and research-based approach conceptualizing the construct of *specialized knowledge for teaching mathematics*. The crucial aspect of the conceptualization is its focus on the *form* of knowledge for teaching, in addition to the content that has been given the most attention in recent years. In more detail, the underlying philosophy of the theoretical framework is the assumption that the defining features of mathematics teacher knowledge cannot be described in terms of more or deeper knowledge bases but in terms of a fundamentally different kind of knowledge, namely the result of a *transformation* of knowledge from various knowledge bases. This transformation perspective implies the view that initial knowledge bases are inextricably combined and restructured into a new form of knowledge that possesses distinct characteristics that were not present in their original form. In the case of *specialized subject matter knowledge for teaching*, for instance, the transformation of intuitive and formal mathematical knowledge is considered as a new form of subject matter knowledge – as specialized pure subject matter knowledge promoting the learning of students. This kind of knowledge is considered as being crucial for teaching mathematics at an upper-secondary level. With this, the theoretical framework takes into account upper-secondary mathematics teachers in contrast to elementary and lower-secondary teachers that have been the predominant focus on past conceptualizations of teacher knowledge. Since all teaching knowledge is contextually bound, the design of the planned conceptual framework focuses on the deep exploration of a few *concept*-specific subcategories of specialized knowledge for teaching taking into account key mathematical concepts in a particular mathematical domain, namely calculus, rather than to cover a great deal of material in a superficial way.

References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14

ARE STUDENTS BETTER AT VALIDATION AFTER AN INQUIRY-BASED TRANSITION-TO-PROOF COURSE?

Annie Selden
New Mexico State University

John Selden
New Mexico State University

This paper presents the results of a study of the proof validation abilities and behaviors of sixteen undergraduates after taking an inquiry-based transition-to-proof course. Students were interviewed individually towards the end of the course using the same protocol that we had used earlier at the beginning of a similar course (*Selden and Selden, 2003). Results include a description of the students' observed validation behaviors, a description of their proffered evaluative comments, and the, perhaps counterintuitive, suggestion that taking an inquiry-based transition-to-proof course does* ***not*** *seem to enhance validation abilities. We also discuss distinctions between proof validation, proof comprehension, proof construction and proof evaluation and the need for research on their interrelation*.

Key words: Transition-to-proof, Proof, Validation

This paper presents the results of a study of the proof validation behaviors of 16 undergraduates after taking an inquiry-based transition-to-proof course emphasizing proof construction. Students were interviewed individually towards the end of the course employing the same protocol used in our earlier study (Selden & Selden, 2003). Here we regard proofs as we did then and use our previous description of *proof validation* as reading and reflection on proofs to determine their correctness (p. 5).

We provide a detailed description of the validation behaviors the 16 undergraduates took – something either not done, or only partially done, in prior validations studies and perhaps not at all for this level of student. Past validation studies include: first-year Irish undergraduates' validations and evaluations (Pfeiffer, 2011); U.S. undergraduates' validations at the beginning of a transition-to-proof course (Selden & Selden, 2003); U.S. mathematics majors' validation practices across several content domains (Ko & Knuth, 2013); U.S. mathematicians' validations (Weber, 2008); and U.K. novices' and experts' reading of proofs to compare their validation behaviors (Inglis & Alcock, 2012).

Our ultimate goal is to understand the process of proof validation. Our specific research question was: Would taking an inquiry-based transition-to-proof course that emphasized proof construction significantly enhance students' proof validation abilities?

Setting: The Course and the Students

The course the participants attended is meant as a 2nd year transition-to-proof course for mathematics majors, but is often taken by a variety of majors and by more advanced students.[1] The students were given notes with definitions, some explanations, and requests for examples and statements of theorems to prove. They proved the theorems outside of class and presented their proofs in class and received extensive critiques. In addition, about once a week, the class worked in groups to co-construct proofs of upcoming theorems in the notes. Sometimes, if the students seemed to need it, there were mini-lectures on topics such as logic. Sixteen of the 17 enrolled students opted to participate in the study for extra credit. Of these, 81% (13 of 16) were either mathematics majors, secondary education mathematics majors, or

[1] We have found that students are often afraid of a transition-to-proof course, and that instead of taking it before courses like abstract algebra and real analysis (with which it is supposed to help), they take it later.

were in mathematics-related fields (e.g., electrical engineering, civil engineering, computer science). In addition to homework, which consisted of attempts to prove the next two or three theorems in the notes, and presenting their attempted proofs in class, the students had mid-term and final examinations. Topics included sets, functions, continuity, and beginning abstract algebra. The teaching aim was to have students experience constructing as many different kinds of proofs as possible, especially in abstract algebra and real analysis.

Methodology: The Conduct of the Interviews

Interviews were conducted outside of class during the final two weeks of the course. The students received extra credit for participating and signed up for one-hour time slots. They were told that they need not study for this extra credit session. The protocol was the same as that of our earlier validation study (Selden & Selden, 2003). Upon arrival, participants were first informed that they were going to validate four student-constructed "proofs" of a single number theory theorem. They were asked to think aloud and to decide whether the purported proofs were indeed proofs. Only "Proof (b)" was. Participants were encouraged to ask clarification questions and informed that the interviewer would decide whether a question could be answered. They were given the same Fact Sheet (see Appendix 1) about multiples of 3 provided to the participants of our earlier study.

There were four phases to the interview: A warm-up phase during which the students gave examples of the theorem: *For any positive integer n, if* n^2 *is a multiple of* 3, *then n is a multiple of* 3 and then tried to prove it; a second phase during which they validated, one-by-one, the four purported (student-constructed) proofs of the theorem; a third phase during which they were able to reconsider the four purported proofs (presented altogether on one sheet of paper), and a fourth debrief phase during which they answered questions about how they normally read proofs. (See Appendix 2 for details.)

The interviews were audio recorded. The participants[2] wrote as much or as little as they wanted on the sheets with the purported proofs. Participants took as much time as they wanted to validate each proof, with one participant initially taking 25 minutes to validate "Proof (a)". The interviewer answered an occasional clarification question, such as the meaning of the vertical bar in $3|n^2$, but otherwise only took notes, and handed the participants the next printed page when they were ready for it. Our 2003 paper contains a detailed textual analysis of the purported proofs (pp. 10-18).

The data collected included: the sheets on which the participants wrote, the interviewer's notes, and the recordings of the interviews. These data were analyzed multiple times to note anything that might be of interest. Tallies were made of such things as: the number of correct judgments made by each participant individually; the percentage of correct judgments made by the participants (as a group) at the end of Phase 2 and again at the end of Phase 3; the validation behaviors that the participants were observed by the interviewer to have taken; the validation comments that the participants proffered; the amount of time taken by each participant to validate each of the purported proofs; the number of times each participant reread each purported proof; the number of participants who underlined or circled parts of the purported proofs; the number of times the participants substituted numbers for *n*; and the number of times the participants consulted the Fact Sheet. Many of these are indicated below.

Observed Participants' Validation Behaviors

All participants appeared to take the task very seriously and some participants spent a

[2] Because the purported proofs were constructed by undergraduate students and because the participants in this study were also undergraduate students, we will henceforth refer to the undergraduates in this study as "participants" to avoid confusion.

great deal of time validating at least one of the purported proofs. For example, LH[3] initially took 25 minutes to validate “Proof (a)” before going on, and VL initially took 20 minutes to validate “Proof (b)”. The minimum, maximum, and mean times for validating each purported proof are given in Table 1.

Table 1: Time (in minutes) taken initially to validate the purported proofs (during Phase 2)

	“Proof (a)”	“Proof (b)”	“Proof (c)”	“Proof (d)”
Maximum time	25	20	16	9
Minimum time	5	2	3	2
Mean time	8.8	8.5	6.3	4.5

The following (probably beneficial) behaviors[4] were observed as being enacted by the participants; the percentages and absolute numbers are given in parentheses:

1. underlined, or circled, parts of the purported proofs (100%, 16);
2. pointed with their pencils or fingers to words or phrases, as they read along linearly (50%, 8);
3. checked the algebra, for example, by “foiling” $(3n+1)^2$ (62.5%, 10);
4. substituted numbers for n to check the purported equalities (37.5%, 6);
5. reread all, or parts of, the purported proofs (87.5%, 14);
6. consulted the Fact Sheet to check something about multiples of 3 (56.25%, 9).

Summarizing the above, participants used focus/reflection aids (1. & 2.); checked computations or tested examples (3. & 4.); revisited important points – perhaps as a protection against “mind wandering” (5.); and checked their own knowledge (6.).

Participants’ Proffered Evaluative Comments

The participants sometimes voiced what they didn’t like about the purported proofs. For example, CY objected to “Proof (b)” being referred to as a proof by contradiction. He insisted it was a contrapositive proof and twice crossed out the final words “we have a proof by contradiction”. Fourteen (87.5%) mentioned the lack of a proof framework,[5] or an equivalent, even though the interviewer had informed them at the outset that the students who wrote the purported proofs had not been taught to construct proof frameworks.

Below are some additional features that seemed to bother some participants:

1. lack of clarity in the way the purported proofs were written, referring to parts of them as “confusing”, “convoluted”, “a mess”, or not “making sense” (68.75%);
2. the notation, which one participant called “wacky”;
3. the fact that “Proof (d)” started with n, then introduced m, and did not go back to n;
4. not knowing what the students who had constructed the purported proofs knew or were allowed to assume:
5. having too much, or too little, information in a purported proof. For example, one participant said there was “not enough evidence for a contradiction” in

[3] Initials, like LH, designate individual participants.
[4] Ko and Knuth (2013, p. 27) referred to validation behaviors, such as checking line-by-line or example-based reasoning as “strategies” for validating proofs. We prefer the term “behaviors” as the act of underlining or circling parts of proofs is evidence of focus, not strategy, which usually entails a plan.
[5] A *proof framework* is a “representation of the ‘top level’ logical structure of a proof, which does not depend on a detailed knowledge of the mathematical concepts, but is rich enough to allow the reconstruction of the statement being proved or one equivalent to it.” (Selden & Selden, 1995, p. 129). In practice, in this transition-to-proof course, this meant writing the hypotheses at the top of the nascent proof, leaving a blank space for the details, and writing the conclusion at the bottom of the proof, and perhaps, also unpacking the conclusion and writing as much as possible of the structure of the proof.

"Proof (b)";
6. the "gap" in "Proof (c)" which was remarked on by six participants.

Individual Participants' Comments

Below are some participants' comments,[6] indicating which are local and which are overall. The overall comments often seemed to have more to do with making sense, having enough information, or being a "strong" proof, rather than with the structure of the purported proofs. Indeed, no participant even commented on the strange division of "Proof (a)" into odd and even cases. This general lack of global, structural comments is similar to prior findings (e.g., Inglis & Alcock, 2012; Selden & Selden, 2003).

Local comments on "Proof (a)": "[I] don't like the string of = s." (MO). "3n+1, if n=1, is not odd, [rather it] would be even." (KW). "This [pointing to $n^2 = 9n^2$] isn't equal." (AF).

Overall comments on "Proof (a)": "[It] needs more explanation -- I can't see where they are going." (CL). "[The] first case doesn't seem right." (CY). "Not going where they need to go." (KW). "Not a proper proof". (FR). "Partial proof". (MO).

Local comments on "Proof (b)": "Not seeing the closing statement." (FR). "Not a proof because we don't introduce n, but we use n." (KK).

Overall comments on "Proof (b)": "[This makes] a lot more sense to me [than "Proof (a)"]" (CL). "[It's] not written well." (SS). "[I] feel like it's a proof because [they're] showing that the two integers in between are not multiples of 3" (AF).

Local comments on "Proof (c)": Commenting on the use of the universal quantifier with *x*, "[The bit about] where *x* is *any* integer worries me" (CJ).

Overall comments on Proof (c)": "Just can't get my head around [it]." (CY). "Need more information. Don't buy it." (CJ). "[This one is] closer [to a proof] than the others." (KK). "Sound proof". (MO).

Local comments on "Proof (d)": "Why would you use *m*? ... [It's] kind of confusing with that *m*." (LH).

Overall comments on "Proof (d)": "[He is] putting [in] more information than needs to be [there]. [This does] not help his proof." (MO). "Not a strong proof." (LH).

We note that the sample participants' comments, given in the above two sections, do not focus just on whether the theorem has been proved. We suspect participants might have had difficulty separating matters of validity from matters of style and personal preference, or even from their own confusion.

What Participants Said They Do When Reading Proofs

In answer to the final debrief questions, all participants said that they check every step of a proof or read a proof line-by-line. All said they reread a proof several times or as many times as needed. All, but one, said that they expand proofs by making calculations or making subproofs. In addition, some volunteered that they work through proofs with an example, write on scratch paper, read aloud, or look for the framework. All of this can be helpful. Furthermore, ten (62.5%) said they tell if a proof is correct by whether it "makes sense" or they "understand it". These are cognitive feelings that, with experience, can be useful. Four (25%) said a proof is incorrect if it has a mistake, and four (25%) said a proof is correct "if they prove what they set out to prove." These last two views of proof call for some caution.

Discussion and Conclusions

In answer to the initial research question, the participants in this study took their task very seriously, but made fewer final correct judgments (73% vs. 81%) than the

[6] After each comment, in parentheses, we have indicated the participant who made the comment.

undergraduates studied earlier (Selden & Selden, 2003) despite, as a group, being somewhat further along academically. In this study, 56% (9 of 16) of the participants were in their 4th year of university, whereas just 37.5% (3 of 8) of the undergraduates in our earlier study were in their 4th year. Because the participants in this study were completing an inquiry-based transition-to-proof course emphasizing proof construction, we conjectured they would be better at proof validation than those at the beginning of our earlier transition-to-proof course (Selden & Selden, 2003), but they weren't. We have tentatively concluded that if one wants undergraduates to learn to validate "messy" student-constructed, purported proofs, in a reliable way, one needs to teach validation explicitly. We stress this because it is counterintuitive. As students most mathematicians have received considerable implicit proof construction instruction through feedback on assessments. However, most have received no validation instruction, but are very skilled at it.

Future Research

In addition to proof validation, there are three additional related concepts in the literature: proof comprehension, proof construction, and proof evaluation. There has been little research on how these four concepts are related. In this study, we investigated one of them -- whether improving undergraduates' proof construction abilities would enhance their proof validation abilities and obtained some negative evidence.

Proof comprehension means understanding a (textbook or lecture) proof. Mejia-Ramos, Fuller, Weber, Rhoads, and Samkoff (2012) have given an assessment model for proof comprehension, and thereby described it in practical terms. Examples of their assessment items include: Write the given statement in your own words. Identify the type of proof framework. Make explicit an implicit warrant in the proof. Provide a summary of the proof.

Proof construction means constructing correct proofs at the level expected of mathematics students (depending which year they are in their program of study).

Proof evaluation was described by Pfeiffer (2011) as "determining whether a proof is correct and establishes the truth of a statement (validation) and also how good it is regarding a wider range of features such as clarity, context, sufficiency without excess, insight, convincingness or enhancement of understanding." (p. 5).

While it is still an open question as to how these four concepts are related, in addition to our study, Pfeiffer (2011) conjectured that practice in proof evaluation could help undergraduates appreciate the role of proofs and also help them in constructing proofs for themselves. She obtained some positive evidence, but her conjecture needs further investigation. As for proof comprehension, it is an open question as to whether practice in comprehension would help any of proof evaluation, proof validation, or proof construction.

In addition, there is anecdotal evidence that some of today's transition-to-proof courses/textbooks are thought to be inadequate for the task of actually transitioning students from lower level undergraduate mathematics courses to upper level undergraduate proof-based courses, such as abstract algebra and real analysis. Whether this is the case, and to what degree, should be investigated.

Finally, we feel that there is a need to develop characteristics of a reasonable learning progression for tertiary proof construction, going from *novice*[7] (lower-division mathematics students) to *competent* (upper-division mathematics students), on to *proficient* (mathematics graduate students), and eventually to *expert* (mathematicians).

[7] The terms *novice*, *proficient*, *competent*, and *expert* have been adapted from the Dreyfus and Dreyfus (1986) novice-to-expert scale of skill acquisition.

References

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the age of the computer*. Oxford: Blackwell.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*(4), 358-390.

Ko, Y.-Y., & Knuth, E. J. (2013). Validating proofs and counterexamples across content domains: Practice of importance for mathematics majors. *Journal of Mathematical Behavior*, *32*, 20-35.

Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics 79*(1), 3-18.

Pfeiffer, K.(2011). Features and purposes of mathematical proofs in the view of novice students: Observations from proof validation and evaluation performances. Doctoral dissertation, National University of Ireland, Galway.

Selden, A., & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem*? Journal for Research in Mathematics Education*, *34*(1), 4-36.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29,* 123-151.

Weber, K. (2008). How do mathematicians determine if an argument is correct? *Journal for Research in Mathematics Education*, *39*(4), 431-439.

Appendix 1: Fact Sheet

[from Selden and Selden (2003), p. 32]

FACT 1. The positive integers, $\mathbf{Z}^+$, can be divided up into three kinds of integers -- those of the form $3n$ for some integer n, those of the form $3n + 1$ for some integer n, and those of the form $3n + 2$ for some integer n.

For example,

1,	2,	3,	4,	5,	6,	7,	8,	9,	10,	11…
		$3n$	$3n+1$	$3n+2$	$3n$	$3n+1$	$3n+2$			
			where $n = 1$			where $n = 2$				

FACT 2. Integers of the form $3n$ (that is, 3, 6, 9, 12, . . .) are called *multiples of* 3.

FACT 3. No integer can be of two of these kinds simultaneously. So m is not a multiple of 3 means the same as m is of the form $3n$+1 or $3n$+2.

Appendix 2: Interview Protocol

[from Selden and Selden (2003), pp. 32-33]

PHASE 1: 'Warm Up' Exercises

For any positive integer n, if n^2 is a multiple of 3, *then n is a multiple of* 3.

1. Explain, in your own words, what the above statement says.
2. Give some examples of the above statement.
3. Does the above statement seem to be true? How do you tell?
4. Do you think you could give a proof of the above statement?

PHASE 2: Sequential consideration of 'Proofs' (a), (b), (c), (d). [The purported proofs were presented to the participants, one page at a time, during this Phase. The purported proofs are given below under Phase 3.]

PHASE 3: 'Recap' on the 'Proofs'

Below are several purported proofs of the following statement:

For any positive integer n, if n^2 is a multiple of 3, *then n is a multiple of* 3.

For each one, decide whether or not it is a proof. Try to "think out loud" so you can let me in on your decision process. If it is not a proof, point out which part(s) are problematic. If you can, say where, or in what ways, the purported proof has gone wrong.

(a). Proof: Assume that n^2 is an odd positive integer that is divisible by 3. That is $n^2 = (3n + 1)^2 = 9n^2 + 6n + 1 = 3n(n + 2) + 1$. Therefore, n^2 is divisible by 3. Assume that n^2 is even and a multiple of 3. That is $n^2 = (3n)^2 = 9n^2 = 3n(3n)$. Therefore, n^2 is a multiple of 3. If we factor $n^2 = 9n^2$, we get $3n(3n)$; which means that n is a multiple of 3. ■

(b). Proof: Suppose to the contrary that n is not a multiple of 3. We will let $3k$ be a positive integer that is a multiple of 3, so that $3k + 1$ and $3k + 2$ are integers that are not multiples of 3. Now $n^2 = (3k + 1)^2 = 9k^2 + 6k + 1 = 3(3k^2 + 2k) + 1$. Since $3(3k^2 + 2k)$ is a multiple of 3, $3(3k^2 + 2k) + 1$ is not. Now we will do the other possibility, $3k + 2$. So, $n^2 = (3k + 2)^2 = 9k^2 + 12k + 4 = 3(3k^2 + 4k + 1) + 1$ is not a multiple of 3. Because n^2 is not a multiple of 3, we have a contradiction. ■

(c). Proof: Let n be an integer such that $n^2 = 3x$ where x is any integer. Then $3|n^2$. Since $n^2 = 3x$, $nn = 3x$. Thus $3|n$. Therefore if n^2 is a multiple of 3, then n is a multiple of 3. ■

(d). Proof: Let n be a positive integer such that n^2 is a multiple of 3. Then $n = 3m$ where $m \in \mathbf{Z}^+$. So $n^2 = (3m)^2 = 9m^2 = 3(3m^2)$. This breaks down into $3m$ times $3m$ which shows that m is a multiple of 3. ■

PHASE 4. Final Questions

1. When you read a proof is there anything different you do, say, than in reading a newspaper?
2. Specifically, what do you do when you read a proof?
3. Do you check every step?
4. Do you read it more than once? How many times?
5. Do you make small subproofs or expand steps?
6. How do you tell when a proof is correct or incorrect?
7. How do you know a proof proves this theorem instead of some other theorem?
8. Why do we have proofs?

118

AN ANALYSIS OF TRANSITION-TO-PROOF COURSE STUDENTS' PROOF CONSTRUCTIONS WITH A VIEW TOWARDS COURSE REDESIGN

Ahmed Benkhalti	John Selden	Annie Selden
New Mexico State	New Mexico State	New Mexico State

The purpose of the study was to gain knowledge about undergraduate transition-to-proof course students' proving difficulties. We analyzed the final examination papers of students in one such course. Our perspective included the sometimes automated links between situations and (mental, as well as physical) actions. We have tentatively identified process, rather than content, categories of difficulties such as nonstandard language/notation, insufficient warrants, and extraneous statements. The ultimate goal is to use an understanding of some of these categories as pedagogical content knowledge with which to redesign an existing transition-to-proof course.

Key words: Transition-to-proof, Proof construction, Pedagogical content knowledge, Actions, Proof framework

This paper presents part of an analysis of transition-to-proof course students' final examinations in an effort to categorize their proving difficulties and develop pedagogical content knowledge contributing to course redesign. Our goal is to go beyond content errors to difficulties of process and cognition and to consider links between situations and actions.

While some studies of students' proving difficulties have been made before, they have not been so closely aimed at course design. Also, several studies have been conducted with students who were mathematically more advanced. For example, Selden and Selden (1987) examined errors and misconceptions in undergraduate abstract algebra students' proof attempts. The difficulties reported there have little in common with those observed in this study. In addition, Weber's (2001) study, contrasting undergraduate abstract algebra students with doctoral students in algebra, showed that the latter had strategic (content) knowledge to use in making algebra proofs that the undergraduates did not have. Our study, in contrast, gives insight into the proving difficulties of relative beginners, that is, undergraduate students at the end of a transition-to-proof course. We note that Moore (1994) observed a traditionally taught transition-to-proof course and reported seven student proving difficulties, some of which do overlap with our categories, but in general, our categories are more fine-grained. In addition, Baker and Campbell (2004) reported three observations of somewhat less sophisticated transition-to-proof course students.

Selden and Selden (1995) did observe process difficulties in unpacking the logic of informal mathematical statements. They reported that informal statements that departed from the simplest natural language rendering of predicate and propositional calculus were difficult for students to unpack and hence difficult to prove. This information was used in designing our current course.

Conduct of the Study

We analyzed all 16 take-home and all 16 in-class final examination papers from one transition-to-proof course at a mid-sized U.S. Ph.D.-granting university. This inquiry-based course was taught entirely from notes with students constructing original proofs and receiving critiques in class. In order to coordinate with later courses, the notes included some theorems about sets, functions, real analysis, and abstract algebra. To maximize student proof presentation, logic was taught in context as the need arose, mainly through the discussion of students' logical errors. The items on the examinations consisted entirely of requests for original proofs of theorems new to the students, but based on the course notes, which were available during the examination.

The papers were analyzed through several iterations, looking for categories of students' proving difficulties until the researchers came to an agreement about categories emerging from the data. The categories were chosen at a level of abstraction above specific mathematical topics so they would reflect *process* difficulties. For example, some were about students not unpacking a conclusion, as opposed to holding a misconception about groups.

We adopted the perspective that a proof construction consists of a sequence of mental or physical actions, some of which do not appear in the final proof. Such actions are driven by understandings of situations in the developing proof construction (Selden, McKee, & Selden, 2010). With repetition, a situation-action link may be automated (Bargh, 2014), yielding behavioral knowledge (Selden & Selden, 2009, p. 343) and thus reducing the call on working memory. One important action is the construction of a proof framework, somewhat similar to that described in Selden and Selden (1995). Writing a proof framework consists not only of writing the hypothesis at the beginning of a proof and the conclusion at its end, but also of doing the same for any subproof, and in particular, for the unpacked conclusion. We have included in our categories beneficial actions that some students did not take, such as writing a full proof framework, as well as detrimental actions that they did take, such as including definitions in proofs. (See 3. Beneficial actions … .)

Categories

In our initial analysis, we allowed categories within categories and hope that their hierarchy will help identify the most important needed interventions. We have thus far tentatively identified the following categories: extraneous statements, inadequate proof framework, unfinished proof, assumption of the negation of a previously established fact, incorrect deduction, nonstandard language/notation, failure to unpack the hypothesis or conclusion, insufficient warrant, assumption of all or part of the conclusion, assertion of an untrue "result", wrong or incorrectly used definitions, difficulties with proof by contradiction, computational errors, misuse of logic, failure to use cases when appropriate, inappropriately mimicking a prior proof, and omitting beneficial actions and taking detrimental ones. Below we illustrate three of these categories:

1. Nonstandard language/notation

In an attempt to prove that the split domain function h, defined by $h(x) = f(x)$ if $x \geq a$ and $h(\mathrm{x}) = g(x)$ if $x < a$, is continuous at a, given that both f and g are continuous at a and $f(a) = g(a)$, one student (4A.3) wrote: "$|f(x)\text{-}f(a)|< \varepsilon/2 - |g(x)\text{-}g(a)|< \varepsilon/2$". This action, subtracting a statement such as "$|g(x)\text{-}g(a)|< \varepsilon/2$", from another statement, violates normal mathematical syntax. Subtraction is an arithmetic operation used between numbers or variable representing numbers, not a logical operation used between statements. How to prevent students from taking such nonsensical actions is an interesting pedagogical question.

2. Unfinished proof

In an attempt to prove that $f(G)$ is a group given that S and T are semigroups, $f{:}S \rightarrow T$ is a homomorphism, and G is a subgroup of S, Student 9B.4 did not warrant or show that $f(G)$ is a semigroup, that is, that $f(G)$ is nonempty and closed under the operation, but did attempt to show the existence of an identity and inverses. (See Appendix.) Student 9B.4 appears not to have known the definition of group, a difficulty noted by Moore (1994), but in fact had it available.

3. Beneficial actions not taken and detrimental actions taken

In an attempt to prove that a semigroup S is commutative, given that for all $s \in S$ we have $ss = e$, where e is an identity of S, Student 5A.4 did not write a full proof framework. In particular, Student 5A.4 did not use the unpacked conclusion, namely that $ab = ba$ for all $a, b \in S$, to structure the proof, that is, Student 5A.4 did not introduce an arbitrary $a, b \in$ S into the proof. This was a beneficial action that Student 5A.4 did not take take. (See Appendix.) We think that had Student 5A.4 written the full proof framework and "explored" the equation $(ab)(ab) = e$ and its consequences, he or she might have been able to construct a correct proof. We also think student, like 5A.4, would benefit from explicit instruction in this sort of "exploration".

Discussion

Focusing on abstraction above the level of specific mathematical topics and on automated actions driven by (inner) interpretations of situations suggests that deductive reasoning is not mainly an interaction of logic and content familiarity, but also depends on several kinds of behavioral and procedural knowledge. In addition to adding a line to the emerging proof or a sketch to scratch work, such behavioral knowledge can yield "meta-actions" (meant to influence one's own thinking) and actions influenced by (cognitive) feelings or unconscious priming (Bargh, 2014). For example Student 5A.4 (Example 3 above) needs to learn when to write a proof framework and not to write things like Lines 3, 4, 5, and 6 (see Appendix). In addition, Student 5A.4 needs to learn when to "explore", that is, create and manipulate objects like $abab = e$ without knowing this will be useful—actions requiring a feeling of self-efficacy (Selden & Selden, to appear).

Discussion Questions

1. How could students be taught to autonomously take better actions than 5A.4 (Example 3 above)?
2. Student 9B.4 (Example 2 above) appears not to know the definition of a group, but he/she had access to it. What should 9B.4 be taught about using definitions?

References

Baker, D., & Campbell, C. (2004). Fostering the development of mathematical thinking: Observations from a proofs course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 14*(4), 345-353.

Bargh, J. A. (2014). Our unconscious mind. *Scientific American, 310*(1), 30-37.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*, 249-266.

Selden, A., McKee, K., & Selden, J. (2010). Affect, behavioural schemas and the proving process, *International Journal of Mathematical Education in Science and Technology, 41*(2), 199-215.

Selden A., & Selden, J. (in press). The roles of behavioral schemas, persistence, and self-efficacy in proof construction. In B. Ubuz, C. Hasar, & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 248-255). Ankara, Turkey: Middle East Technical University.

Selden A., & Selden, J. (1987). Errors and misconceptions in college level theorem proving. In J. D. Novak (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics* Vol. III (pp. 457-470), Ithaca, NY: Cornell University.

Selden, J., & Selden, A. (2009). Teaching proving by coordinating aspects of proofs with students' abilities. In D.A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and Learning Proof Across Grades: A K-16 Perspective* (pp. 339-354). New York/Washington, DC: Routledge/National Council of Teachers of Mathematics.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*, 123-151.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge, *Educational Studies in Mathematics, 48*, 101-119.

Appendix

Below we give two student-constructed "proofs" together with sample correct proofs. We note that student proving difficulties seem often not to be isolated, but occur in combinations.

Theorem. Let S be a semigroup with an identity element e. If, for all s in S, $ss = e$, then S is commutative.

Student "Proof" 5A.4

Proof: Let S be a semigroup with an identiy element, e. Let $s \in S$ such that $ss = e$.
Because e is an identity element, $es = se = s$.
Now, $s = se = s(ss)$.
Since S is a semigroup, $(ss)s = es = s$.
Thus $es = se$.
Therefore, S is commutative. QED.

Sample Correct Proof

Proof: Let S be a semigroup with identity e.
Suppose for all $s \in S$, $ss = e$.
Let a, b be elements in S.
Now $abab = e$, so $(abab)b = eb = b$.
But $(abab)b = aba(bb) = (aba)e = aba$.
Thus $aba = b$, so, $(aba)a = ba$, and $(aba)a = ab(aa) = abe = ab$.
Thus $ba = ab$.
Therefore, S is commutative. QED

SCRATCHWORK:
7.1: A semigroup is called commutative or Abelian if, for each a and $b \in S$, $ab = ba$.
7.5: An element e of a semigroup S is called an identity element of S if, for all $s \in S$, $es = se = s$.

Comment: Student 5A.4's Line 2 probably meant "Suppose, for all $s \in S$, $ss = e$." Line 3 violates the mathematical norm of not including in a proof definitions that can easily be found outside the proof. Also Lines 4, 5, and 6 could not contribute to the proof because they do not involve two variables (necessary to show commutativity). While Lines 3, 4, 5, and 6 are true, we conjecture they should not have been included for psychological reasons because they might have wrongly suggested to Student 5A.4 that something useful had been done.

Theorem. Let S and T be semigroups and $f{:}S\to T$ be a homomorphism. If G is a subset of S and G is a group with identity e, then $f(G)$ is a group.

Student "Proof" 9B.4

Proof: Let S and T be semigroups and $f{:}S\to T$ be a homomorphism. Suppose $G \subseteq S$ and G is a group with identity e. Since G is a group and it has identity e, then for each element g in G there is an element g' in G such that $gg' = g'g = e$. Since f is a homomorphism, then for each element $x \in S$ and $y \in S$, $f(xy)=f(x)f(y)$. Since $G \subseteq S$, then $f(gg')=f(g)f(g')$. So $f(gg') = f(g'g) = f(e)$. So $f(G)$ has a an element $f(e)$ since f is a function. Therefore, $f(G)$ is a group. QED.

Sample Correct Proof

Proof: Let S and T be semigroups and $f{:}S\to T$ be a homomorphism. Let G be a subset of S and G be a group with identity e.

Part 1.Note that G is a subsemigroup of S so, by Theorem 20.4, $f(G)$ is a semigroup.

Part 2. Let $y \in f(G)$. Then there is $x \in G$ so that $f(x) = y$. Now $f(e) \in f(G)$ and $f(e)y = f(e)\, f(x) = f(ex) = f(x) = y$. Similarly, $y\, f(e) = y$. Thus $f(e)$ is an identity for $f(G)$.

Part 3. Let q in $f(G)$. Then there is $p \in G$ so that $f(p) = q$. Now because G is a group, there is $p' \in G$ so that $pp' = p'p = e$. Thus $q\, f(p') = f(p)\, f(p') = f(pp') = f(e)$, and $f(p')q = f(p')\, f(p) = f(p'\, p) = f(e)$. Thus, each $q \in f(G)$ has an inverse, $f(p')$, in $f(G)$.

Therefore, is a $f(G)$ group. QED

Comment: After writing the beginning and end of the proof (which could be considered the start of a proof framework), Student 9B.4 should have continued constructing the framework by unpacking and using the meaning of $f(G)$ is a group. This has three parts and is about $f(G)$, not G. Including the three parts could have further "structured" the proof. Instead, Student 9B.4 wrote into the proof the definition of G being a group and f being a homomorphism. These were not useful because the conclusion is about $f(G)$, not G. We conjecture that those two actions served only to wrongly suggest to Student 9B.4 that progress on a proof had been made.

119

STUDENTS' UNDERSTANDING OF EXPONENTIAL FUNCTIONS IN THE CONTEXT OF FINANCIAL MATHEMATICS

Natalie E. Selinski
San Diego State University

Exponential functions are one of the most critical mathematical topics used by students in financial mathematics. This presentation explores university finance students' notion of exponential function from two sets of data. First, I use data collected through surveys to examine students' understanding of exponential function in general and, more specifically, to identify the extent to which students conflate exponential functions with polynomials. I then draw on data collected in an inquiry-based instructional sequence aimed at improving financial mathematics students' understanding of exponential functions. Results include delineation of what ways of understanding exponential functions are critical to studying financial mathematics and insights into how best to guide students in developing these understandings within the context of their field of study.

Key words: Exponential Functions, Realistic Mathematics Education, Financial Mathematics

Exponential functions are one of the most critical mathematical topics used by students in financial mathematics, but they are also one of the most challenging topics. There exists a wide but scattered body of literature addressing this challenge. Some document the power and limitation of various conceptions and misconceptions of exponential functions (cf. Confrey, 1990; Strom, 2012). Radley (2004) questions the language of exponential functions, suggesting that calling a function with an independent variable in the exponent an exponential function while calling a function with an independent variable in the base a power function (as opposed to a base function) clouds students' understanding of a base and power. Weber (2002) suggests students' limited knowledge about exponential and logarithmic functions stems from weak reasoning about the process of exponentiation.

Other literature provides ideas for alternative instruction approaches. Confrey and Smith (1995) develop an approach to exponential functions using a combination of covariation of functions and an isomorphism between multiplication as repeated addition and exponentiation as repeated multiplication. Webb, van der Kooij, and Geist's (2011) use Realistic Mathematics Education design principles to create an instructional sequence aimed at developing student understanding of logarithms, including a task for approaching exponential growth and exponential functions by contrasting linear and exponential growth.

In this report I build from and extend both these bodies of literature, first by identifying the student conceptions of exponential functions that business students bring to their first semester studies in financial mathematics, then by analyzing the ways in which an instructional sequence similar to that developed by Webb et al. could be used to deepen students' understanding of exponential function as needed in their financial mathematics studies.

Methodology: Selection of Participants and Method of Data Collection and Analysis

Data for this study were collected from students in a remedial mathematics bridging course that ran parallel to the students typical first semester studies in financial mathematics and business law at an applied science university in Germany. All students in the required financial mathematics course had been recommended to attend the bridging course following overall poor performance on a mathematics entrance exam at the beginning of the semester. Nevertheless, participation in the bridging course was voluntary. As such, only 16 out of 40 first semester students attended the bridging course on one or more occasions.

Prior to any discussion of exponents or exponential functions in the bridging course, a survey was administered to establish a baseline of students' understandings of exponents, exponential functions, logarithms, and logarithmic functions. For the purpose of this report, only those questions regarding exponents and exponential functions will be discussed. The following figure gives the relevant survey questions:

1. What is an exponent? Use examples, graphs, formulas, words, definitions, etc. to describe your understanding of exponents.
2. What is an exponential function? Again, use examples, graphs, formulas, words, definitions, etc. to describe your understanding of exponential functions.
3. How are exponents and exponential functions used in financial mathematics?

Figure 1. Survey questions regarding exponents and exponential functions.

The survey was administered in middle of the semester, after the required financial mathematics course had covered cost-, revenue-, and profit-functions as well as interest calculations. On the day the surveys were completed in the bridging course, 12 students were in attendance. All students voluntarily participated in the surveys, which took approximately 10-15 minutes.

In the classes following the survey, two related inquiry-based tasks were used to explore exponential functions in the context of financial mathematics. The introductory parts of these tasks are presented in Figures 2 and 3.

You need 100€ more to buy the latest iPhone, so you borrow it from one of your siblings. Your sister is willing to lend you the money. She would charge you 5.50€ interest for every month you haven't paid her back. How much would you owe your sister after 1 month? 2 months? 3? 4? 5? 12? Write a formula for how much you would owe your sister after t months and draw the graph. Your brother would charge you 5% interest for every month you haven't paid him back. How much would you owe your brother after 1 month? 2 months? 3? 4? 5? 12? Write a formula for how much you would owe your brother after t months and draw the graph.

Figure 2. Introduction to first task: Borrowed Money 1

You still need 100€ to buy the latest iPhone. Unfortunately, when you borrowed money in the past, you rarely paid it back. So there are very few people who will lend you money.

1. Your brother will lend you the money with a nominal interest rate of 100% compounded yearly. How much do you owe your brother after 1 year? 2 years? 3 years? ½ year? Write a formula for this situation.
2. Your cousin will also lend you the money, but you must pay it back in two years. You think this will not be a problem, since in two years you plan to have a good job in finance and should be able to pay him back. Your cousin hasn't decided on a nominal interest rate yet, but will compound the interest yearly. How much will you owe your cousin if the nominal interest rate is 5%? 10%? 50%? 100%? 150%? 200%? Write a formula for this situation.

Figure 3. Introduction to second task: Borrowed Money 2

Particular care in the design of the tasks was given to differentiating between exponential functions and polynomials. Data from these tasks included classroom video-recordings, detailed observational notes taken immediately after class by the instructor, and copies of all written work by students and instructor.

As the semester progressed, attendance in the bridging course decreased and became more sporadic. As such, only two students participated in all three consecutive classes in which exponential functions were discussed and hence only these two students are used as the case study analysis. Case study data will be analyzed using a coding developed out of the survey and classroom data, vetted by other researchers, in the spirit of grounded theory. The framework for students' mathematical noticing as developed by Lobato, Hohensee, and Rhodehamel (2013) will be used to guide the analysis

Preliminary Results

In the survey students identified any function with an exponent as an exponential function. Figure 4 show the responses of two students that represent typical responses from 10 out of 12 students to the second survey question: What is an exponential function?

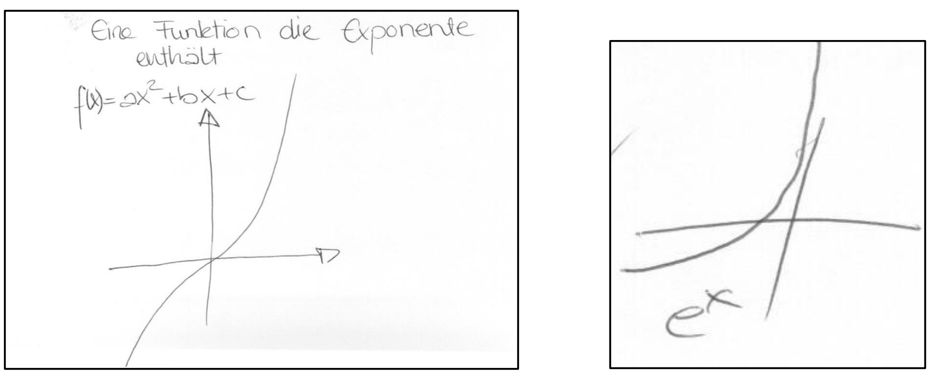


Figure 4. Students describe an exponential function as "a function that contains exponents" [translation], meaning both functions like a quadratic function (left) or the *e* function (right)

In perceiving both polynomials and functions with a known basis and unknown exponent as exponential functions, 8 out of 12 students incorrectly identified applications of exponential functions in financial mathematics, as in the student response to the third survey question (How are exponents and exponential functions used in financial mathematics?) in Figure 5.

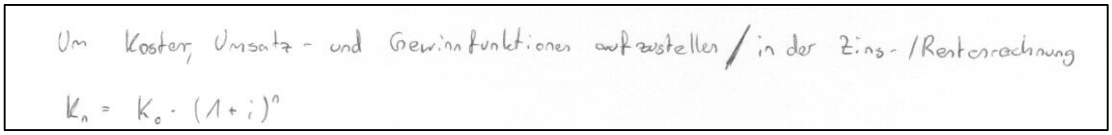


Figure 5. Student response translates to "to create cost-, revenue-, and profit-functions / in interest-/retirement calculations"

Here the student interprets the generalized formula for yearly compounded interest $K_n = K_0 \cdot (1 + i)^n$ as an example of exponential function. However, this formula does not yield an exponential function when the initial principle and time of the investment are parameters and the interest rate is the exponential function. To what extent do students conflate or leave the potential to conflate polynomials and exponential functions? How would this conflation or incorrect understanding of exponential function impact students' studies in financial mathematics?

To begin unpacking these questions, the instructor developed and guided students through the two tasks introduced in Figures 2 and 3, Borrowed Money 1 and 2. In the report, I will present examples from the two students in the case study and their changing understanding of exponential function as evidenced by their work on the tasks Borrowed Money 1 and 2. Examples of this changing understanding of exponential functions, first as a distinction between linear and exponential functions and later as a distinction between polynomials and exponential functions, are shown in Figures 6 and 7, which come from whole class summation of the Borrowed Money tasks.

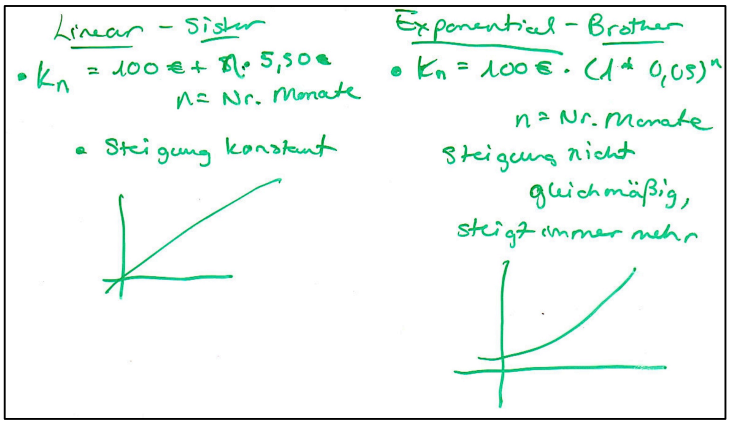


Figure 6. Class summation of the sister's situation versus the brother's situation in Borrowed Money 1. Translation: "Linear – Sister: $K_n = 100€ + n \cdot 5{,}50€, n =$ number of months, constant slope" and "Exponential – Brother: $K_n = 100€ \cdot (1 + 0.05)^n, n =$number of months, slope is not regular, increases always more"

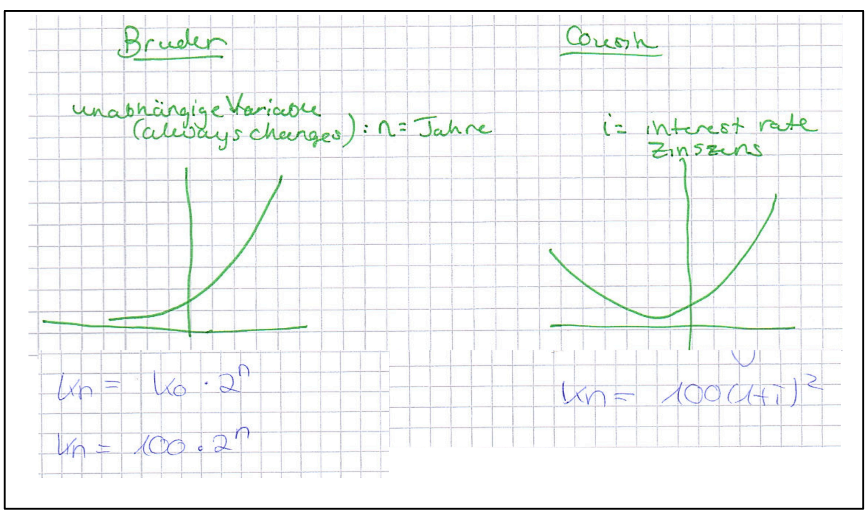


Figure 7. Portion of the class summation of the brother's situation versus the cousin's situation in Borrowed Money 2. Translation: "Brother: independent variable (always changes), *n* = years" and "Cousin: *i* = interest rate"

Analysis of class discussion that produced these figures suggests that students are beginning to see a distinction in these various situations, both in the experientially-real interpretation of the situations and the mathematical behavior. I claim these distinctions can be leveraged in the progress of the instructional sequences to give students reason to draw a distinction between polynomials and exponential functions, both in the general mathematical sense and in the context of financial mathematics.

Questions for Discussion

Existing literature (and for many, personal instructional experience) suggest multiple understandings of exponents and exponential functions beyond what is seen in the student surveys and explored in the presented tasks. What interpretations and ideas regarding exponential functions are necessary to students studying applied fields, particularly financial mathematics? What further tasks or techniques could be used to address students' conflation of exponential functions and polynomials?

As was apparent in the methodology, attendance in these voluntary remedial bridging courses is not consistent enough to develop any lectures or tasks that build off of each other from one class to the next. The only way to ensure more consistent attendance is to use these tasks in the required financial mathematics coursework. To what extent is it possible to incorporate these tasks in a required financial mathematics course? What changes would need to be made to the tasks to push these beyond the remedial level? What changes would occur to the hypothetical learning trajectory when incorporating these tasks in the required financial mathematics course?

References

Confrey, J. (1990). A review of the research on student conceptions in mathematics, science and programming. In C. Cazden (Ed.), *Review of Research in Education 16* (pp. 3-56).

Confrey, J., & Smith, E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education, 26*, 66-86.

Lobato, J., Hohensee, C., & Rhodehamel, B. (2013). Students' mathematical noticing. *Journal for Research in Mathematics Education, 44(5)*, 809-850.

Radley, M. (2004). An exponential function: Is its description not problematic? *The Mathematics Education into the 21st Century Project.* Ciechocinek, Poland. Retreived from *math.unipa.it/~grim/21_project/CiechMahlobo.pdf.*

Strom, A. D. (2009). A case study of a secondary mathematics teacher's understanding of exponential function: A theoretical framework. *The Thirty-fifth American Mathematical Association of Two-Year Colleges Annual Conference.* Las Vegas, NV. Retrieved from *www.amatyc.org/resource/resmgr/2009_conference_proceedings/strom1.ppt*

Webb, D. C., van der Kooij, H., & Geist, M. (2011). Design research in the Netherlands: Introducing logarithms using Realistic Mathematics Education. *Journal for Mathematics Education at Teachers College. 20(1).* 47-52.

Weber, K. (2002). Students' understanding of exponential and logarithmic functions. *Second International Conference on the Teaching of Mathematics.* Crete, Greece: University of Crete.

120

APPROXIMATION: A CONNECTING CONSTRUCT OF THE FIRST-YEAR CALCULUS?

Kimberly S. Sofronas
Emmanuel College

Thomas C. DeFranco
University of Connecticut

Hariharan Swaminathan
University of Connecticut

Nicholas Gorgievski
Nichols College

Charles Vinsonhaler
University of Connecticut

Brianna Wiseman
Emmanuel College

Sam Escolas
Emmanuel College

Abstract

This report will present preliminary findings from a research study designed to investigate calculus instructors' perceptions of approximation as a central concept and possible unifying theme of the first-year calculus. The study will also examine the role approximation plays in participants' self-reported instructional practices. A survey was administered through Qualtrics to a stratified random sample of 3930 mathematicians at higher education institutions throughout the United States with a desired N = 300. Quantitative and qualitative methods were used to analyze the data gathered. Findings from this research will contribute to what is known about the perceptions and teaching practices of calculus instructors regarding the role of approximation in first-year calculus courses. Research-based findings related to the role of the approximation concept in the first-year calculus could have implications for first-year calculus curricula.

Keywords: Calculus, Approximation, Central Concept, Unifying Thread, Curriculum, Perceptions, Higher Education

Significance of the Research Issue

This research originated from findings of a previous study conducted by Sofronas, DeFranco, Vinsonhaler, Gorgievski, Schroeder, and Hamlin (2011) on calculus experts' perceptions of what it means to understand the first-year calculus. In that study, *approximation* was identified as a central concept of the first-year calculus by a third of the 24 calculus experts who participated. This finding raises important questions: Do calculus instructors in the higher education community perceive approximation to be a central idea of the first-year calculus? If so, do calculus instructors in higher education identify approximation as a unifying thread of the first-year calculus? What role does approximation assume in the instructional practices of mathematicians teaching first-year calculus courses in higher education? More research is needed to gain foundational insights into the current views and instructional practices – related to approximation concepts – of mathematicians teaching first-year calculus courses in higher education.

Background Literature

Approximation ideas are often relevant to studies that explore students' understanding of first-year calculus concepts such as function, continuity, limit, derivative, integral, infinite series and more (Asiala et al., 1997; Oehrtman, 2008, 2009; Yang & Gordon, 2008). For example, an exploratory study conducted by Martin (2013) examined differences in the ways in which calculus students and faculty (or graduate student) experts conceptualized Taylor series convergence. In that study, approximation concepts were relevant to task-based interview questions such as those that asked participants from both groups to share their thinking about the purpose of studying Taylor series, approximating functions using Taylor series, and methods for improving those approximations.

Oehrtman (2008, 2009) investigated calculus students' understanding of limit concepts and found that students most commonly reason about limits in terms of approximations. Moreover, he noted that the approximation metaphor for limits most closely resembled the correct mathematical structure underlying the limit. Oehrtman's (2008, 2009) research showed that students were able to use approximation ideas to facilitate the development of a conceptual structure of understanding that would ultimately provide students with the proper foundation for a more formal understanding of limits, if desired.

Gordon (2012) suggested that exposing students in the first-year calculus to ideas from numerical analysis can lead to "…different perspectives and deeper insight into topics that they do see in freshman calculus" (p. 437). According to Gordon (2011), students benefit from examining the errors associated with left- and right-hand Riemann sums to approximate a definite integral whose value is known exactly. Using data analysis techniques to examine patterns in the errors, students derive the Trapezoid Rule, Midpoint Rule and Simpson's Rule and also compare the effectiveness of those methods in approximating a definite integral. The power of the numerical methods for approximating an area under a curve becomes evident when students are then presented with examples such as "… $\int_0^1 e^{-x^2}\,dx$ that cannot be evaluated in closed form by any of the standard integration techniques usually developed in introductory calculus" (Gordon, 2011, p. 149). Nonetheless, Gordon (2012) notes that many calculus instructors shy away from introducing numerical methods in the first-year calculus either because they are not acquainted or not comfortable with them.

There has been some literature support for approximation as a unifying thread of the elementary calculus sequence (Gordon, 2011; Gordon, 2012; Hathaway, 2008; Knisley, 1997; Roberts, 1998; Sofronas et al., 2011; Yang & Gordon, 2008). Hilbert, Schwartz, Seltzer, Maceli, & Robinson (2010) defined *unifying threads* as themes that "…are woven throughout the course, and serve to bind it together into a unified whole" (p. xiii). As cited in Sofronas et al. (2011), "fragmented learning is a major problem in undergraduate mathematics courses (Baroody et al., 2007; Berry & Nyman, 2003; Galbraith & Haines, 2000; Hiebert & Lefevre, 1987; Kannemeyer, 2005). Students form part of the big picture of calculus when they have opportunities to make connections between concepts and can identify the elements of conceptual knowledge that serve as the underlying principles of related procedures (Berry & Nyman, 2003)" (p. 146). Framing calculus curricula around unifying themes, or threads, is one means for promoting a connected understanding of the discipline. Dray and Manogue (2010), for example, argued that differentials can be used in first-year calculus courses to provide a unifying theme that may allow students to build a coherent conception of the calculus.

Roberts (1998) reported on a reform project involving a working group of faculty from 26 liberal arts colleges, which sought to develop a curricular core for a one-year single variable calculus course. While the fundamental theorem easily surfaced as a connecting construct of the first-semester calculus, it was more challenging for the working group to identify a unifying theme for the second-semester calculus. Ultimately, they agreed to "…build the course around 'precision and approximation,' to investigate methodologies that produce exact solutions and when [those] approaches fail, to find ways to obtain approximate solutions with upper bounds on errors" (Roberts, 1998, p. 38). The working group believed, particularly in the case of the second-semester calculus, that approximation was a theme that provided cohesion to the collection of calculus ideas and techniques commonly found in Calculus II courses while – at the same time – "…emphasizing the importance of making approximations" (Roberts, 1998, p. 38).

Likewise, Hathaway (2008) described an approach that used approximation as a unifying theme in the exploration of problems fundamental to the elementary calculus. As they engaged in problem solving, students were encouraged to organize their thinking around the acronym *CAL*: Capitulate, Approximate, and Limit (take). Students must first **capitulate**, or surrender, to the idea that calculus is needed to solve a given problem. In other words, it "…cannot be solved using existing simple algebraic techniques" (p. 543). Students then find a good **approximate** solution. Finally, students take a **limit** of that approximation in an effort to make their approximate solution exact.

At present, no study has comprehensively examined the instructional practices and perceptions related to approximation concepts among mathematicians teaching first-year calculus courses. We hope that the preliminary findings of this study will invigorate discussion about approximation as a possible unifying thread in the first-year calculus, which could have implications for the teaching and learning of first-year calculus courses.

Methodology

Sample

A stratified random sampling method was used to identify the sample for this study. The National Center for Educational Statistics database (http://nces.ed.gov/collegenavigator/) was used to randomly select five higher education institutions - four 4-year and one 2-year institution - from each state in the U.S., as well as Washington D.C., Puerto Rico, and the Virgin Islands. The contact information (i.e., name, institution, and email address) of all mathematicians from those randomly selected institutions was obtained by visiting the institutional websites. A total of 259 institutions were randomly selected through the sampling design. Of those, 77 were excluded from the sample for the following reasons: (a) calculus courses were not offered at the institution; (b) there was no mathematics department at the institution (special-focus institution); or (c) mathematics faculty contact information was not publically available on the institutional website. From the remaining 182 institutions in the sample, a database of 3,930 mathematicians was compiled and all were recruited to participate in an online survey developed for the purpose of this study. Of the 3,930 mathematicians recruited, only those who had taught a first-year calculus course were eligible to participate in the study. A total of N = 279 mathematicians participated in the study.

Survey Instrument

For the purpose of this study, a survey instrument was developed to examine the perceptions and self-reported instructional practices - as they relate to the topic of approximation - of mathematicians who have taught first-year calculus courses in higher education. A review of the literature was conducted to establish item stems for the survey. Content validity was established through consultations with 6 experts in the field. Items stems were added, omitted and refined based upon the feedback of those experts. The survey includes a series of demographic questions, 20 Likert-scale item stems (see Figure 1), an open text box following each Likert-scale item stem to allow participants the option of explaining their rating on the item stem, and two open-response questions. Qualtrics – a secure, internet-based survey technology provider - was used as the platform to create and distribute the survey.

Figure 1. A sample of the Likert-scale item stems (1 = Strongly Disagree, 5 = Strongly Agree) included in the survey instrument.

- In my own teaching of the first-year calculus, I use approximation as a unifying thread to connect many of the key ideas in the calculus curriculum.
- In my own teaching of the first-year calculus, I stress the importance of knowing how good an approximation is.
- In my own teaching of the first-year calculus, I discuss methods for calculating or estimating the error in an approximation.
- In my own teaching of the first-year calculus, I discuss the notion of acceptable levels of error in an approximation.
- In my own teaching of the first-year calculus, I show how the slope of the tangent line can be approximated by slopes of secant lines.

Procedure

Data was collected via Qualtrics for a period of 6 months. Quantitative data was exported to SPSS and analyzed using descriptive statistics and t-tests / analysis of variance procedures for statistically comparing the means of the demographic groups of interest. Qualitative data was coded using a posteriori categorical content analysis techniques. Members of the research team and trained research assistants isolated dominating themes and defined ranges of themes, indicators for the occurrence of a theme and rules applied to the process of coding (Kortendick & Fischer, 1996).

Conclusions and Implications

Preliminary findings of this study will be presented at the RUME-17 conference and will shed light on approximation ideas first-year calculus instructors report emphasizing in their teaching. Differences between demographic groups will be highlighted. Initial findings will also report themes and patterns in calculus instructors' perceptions of approximation both as a central concept and a unifying theme of the first-year calculus curriculum. The alignment between self-reported instructional practices and perceptions of approximation will also be discussed.

References

Asiala, M., Cottrill, J., Dubinsky, E., & Schwingendorf, K. E. (1997). The development of students' graphical understanding of the derivative. *Journal of Mathematical Behavior, 16(4)*, 399–431.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38(2)*, 115–131.

Berry, J. S., & Nyman, M. A. (2003). Promoting students' graphical understanding of the calculus. Journal of Mathematical Behavior, 22, 481–497.

Dray, T. & Manogue, C. (2010). Putting differentials back into calculus. *College Math Journal, 41(2)*, 90 – 100.

Galbraith, P., & Haines, C. (2000). Conceptual mis(understandings) of beginning undergraduates. International Journal of Mathematical Education in Science and Technology, 31(5), 651–678.

Gordon, S. P. (2011). A classroom note on: Building on errors in numerical integration. *Mathematics and Computer Education, 45(2)*, 148 – 156.

Gordon, S. P. (2012). Some surprising errors in numerical differentiation. *PRIMUS, 22(6)*, 437 – 450.

Hathaway, R. J. (2008). A simple acronym for doing calculus : CAL. PRIMUS, 18(6), 542 – 545.

Hiebert, J., & Lefevre, P. (1987). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 1–27). Hillsdale, NJ: Erlbaum.

Hilbert, Schwartz, Seltzer, Maceli, & Robinson (2010). *Calculus: An active approach with projects.* MAA

Kannemeyer, L. (2005). Reference framework for describing and assessing students' understanding in the first year calculus. *International Journal of Mathematical Education in Science and Technology, 36(2–3)*, 271–287.

Knisely, J. (1997). Calculus: A modern perspective. *The American Mathematical Monthly, 104(8)*, 724 – 727.

Kortendick, O. & Fischer, M. (1996). Comparative classifications of ethnographic data:

Constructing the outline of cultural materials. Studies in Anthropology, 11, Center for Social Anthropology and Computing.

Martin, J. (2013). Differences between experts' and students' conceptual images of the mathematical structure of Taylor series convergence. *Educational Studies in Mathematics, 82*, 267 – 283.

Oehrtman, M. (2008). Layers of abstraction: Theory and design for the instruction of limit concepts. In M. P. Carlson & C. Rasmussen (eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics* (pp. 65 – 80). MAA.

Oehrtman, M. (2009). Collapsing dimensions, physical limitation, and other student metaphors for limit concepts. *Journal for Research in Mathematics Education, 40(4)*, 427–459.

Roberts, W. (1998). Calculus in the core. In J.A. Dossey, ed., *Confronting the Core Curriculum: Considering Change in the Undergraduate Mathematics Major*, MAA Notes **45**, MAA, 1998.

Sofronas, K. S., DeFranco, T. C., Vinsonhaler, C., Gorgievski, N., Schroeder, L., & Hamelin, C. (2011). What does it mean to understand the first-year calculus? Perspectives of 24 experts. *Journal of Mathematical Behavior*, 30(2), 131 - 148.

Yang, Y., & Gordon, S. (2008). Finding the best linear approximation to a function. *Mathematics and Computer Education, 42(2)*, 95 – 108.

GEOMETRIC REASONING ON THE COMPLEX PLANE

Hortensia Soto-Johnson and Jonathan Troup

University of Northern Colorado

Using Bakker and Hoffman's (2005) framework on diagrammatic reasoning, we analyzed a video-taped interview to explore two undergraduates' ability to reason geometrically about tasks related to complex variables. Our findings indicate that in order to provide a geometric interpretation, our participants needed to first perform algebraic computations that reduced the current task to a task they knew how to interpret geometrically. These computations appeared to provide them with the pieces required to construct a diagram. Once these pieces were in place the participants used dynamic gesture to enact their geometric interpretations with the aid of their diagram. It appeared that their dynamic gestures assisted with embodying geometric interpretations and as such one particular task was influential throughout the interview. Furthermore, the participants integrated less dynamic gesture as they progressed with similar tasks.

Key words: Algebraic reasoning, Complex variables, Diagrammatic reasoning, Geometric reasoning, Gestures

Introduction and Literature Review

Inspired by Presmeg's (2006) list of potential research questions related to visualization and gesture, we explored undergraduates' geometric reasoning of algebraic equations related to complex variables. Our first research question was: What is the nature of undergraduates' algebraic and geometric reasoning of complex-valued equations and the interplay between the two representations? The second research question was: What is the nature of students' integration of diagrams and gesture as part of their geometric reasoning? Our findings indicate that in order to provide a geometric interpretation, our participants needed to perform algebraic computations to reduce the current task to a task they knew how to interpret geometrically. These computations appeared to provide them with the pieces required to construct a diagram. Once these pieces were in place the participants used dynamic gesture to enact their geometric interpretations with the aid of their diagram. It appeared that their dynamic gestures assisted with embodying geometric interpretations and as such the participants integrated less dynamic gesture as they progressed with similar tasks.

In the past decade, empirical studies centered on students' algebraic and geometric interpretations of complex numbers and functions have begun to emerge (Conner, Rasmussen, Zandieh, & Smith, 2007; Danenhower, 2000, 2006; Nemirovsky, Rasmussen, Sweeney, & Wawro, 2012; Panaoura Elia, Gagatsis, & Giatlilis, 2006). Most of these studies illustrate that students are able to translate between algebraic and geometric representations, but they tend not to utilize the two approaches in tandem. For example, Danenhower (2006) examined undergraduates' ability to convert instantiations of the fraction $\frac{a+ib}{c+id}$ to either Cartesian ($x + iy$) or exponential form ($re^{i\theta}$). Although the undergraduates worked flexibly with the Cartesian form, this was not the case with the exponential form due to the trigonometry involved with the tasks. Furthermore, the undergraduates did not attend to geometric interpretations of the complex number or recognize which form would be more appropriate, both of which could have alleviated much of the computational effort.

In a similar study, Panaoura et al. (2006) explored high school students' ability to navigate between geometric and algebraic representations of complex-valued equations and inequalities. The students tended to be more successful in their responses when given a geometric figure, but not as successful with similar problem-solving tasks. This may indicate "a lack of flexibility in using the geometric approach effectively" (p.700), which could further suggest that students are compartmentalizing symbolic/algebraic, geometric, and verbal representations. In a more recent study, Nemirovsky et al. (2012) provided results that may foster a unified view of algebraic and geometric representations of complex number arithmetic using perceptuo-motor activities. The authors conducted a teaching experiment with preservice secondary teachers, where the students discovered geometric interpretations for adding and multiplying complex numbers. Unlike Danenhower's (2006) and Panaoura's studies, the authors found that their students noticed when their algebraic computations and embodied reasoning disagreed. This research suggests that embodied reasoning of a mathematical procedure with physical models prompted the participants to modify their views of the represented mathematical concepts.

In addition to embodied reasoning and external representations such as Nemirovsky's embodied complex plane, diagrams and gesture can also be utilized as external representations to help students connect algebraic and geometric approaches. Bakker and Hoffman (2005) defined a *diagram* as "a complex sign, which includes icons, indices, and symbols." (p. 339) and represents relationships between mathematical objects. The three steps of diagrammatic reasoning are *diagrammatization*, *experimentation*, and *observation* with *reflection*. In the diagrammatization stage one constructs a diagram or diagrams deemed important or necessary, using an appropriate representational system. Châtelet's (2000) and de Freitas and Sinclair (2012) both note that diagrams and gestures are intimately connected. Diagrams can represent a gesture, and new gestures can be produced as a result of consideration of the diagram. It is thus more natural to view diagrams and gestures in tandem rather than separately, and could therefore further aid students in learning how to connect different representations. In our research, the complex plane served as the primary representational system. The experimentation phase involves recognizing operations that can be done on the diagram. The allowable actions on the diagram are determined by the rules of the system in which one is working. The third step entails observing and reflecting on the results obtained through the diagram(s) and the actions performed on the diagram. This last step could result in abandoning the diagram or creating a new diagram.

Theoretical Perspective

We adopted embodied cognition as our theoretical lens to allow us to explore the connections participants formed between algebraic and geometric representations of complex numbers. As Nemirovsky et al. (2012) demonstrated, this lens allows gesture to serve as a bridge between these two representations. This bridge is natural and intuitive, as Goldin-Meadow (2003) notes that speech and gesture form a single cohesive system of communication and thought, and Sfard (1991) suggests that speech could be more suited to the expression of algebraic ideas and gesture can more easily demonstrate geometric concepts. The perspective of embodied cognition thus facilitates research settings designed to explore participants' connections between the visual and the verbal, between gesture and speech, or between geometry and algebra. Embodied cognition at its core suggests that an organism's personal experience within the physical environment is the primary means by which the organism acquires knowledge. This means that thought is essentially formed by perceived experience with the physical environment. Since embodied cognition relates all thought back to some aspect of the surrounding environment, it is natural to include gesture as a data source within this lens, as it seems to be an expression of thought (Goldin-Meadow, 2003) that actually takes place within the physical environment, perhaps produced as a way of reducing cognitive load (Alibali & Nathan, 2012).

Research Methodology

As part of an undergraduate complex variables course, students completed *GeoGebra* labs to discover the geometric interpretations of the four arithmetic operations and conjugation of complex numbers. After completing these labs, four students participated in video-recorded interviews in pairs of two, where they provided geometric interpretations and made connections to the algebraic representations of 19 facts borrowed from Needham's text, *Visual Complex Analysis* (1997, p.7). The participants shared ideas with one another, questioned one another, and convinced one another of their interpretations; the interviewer probed as needed. Our analysis entailed watching the video to determine which tasks provided the most relevant content related to our research questions. Using *Elan*, we transcribed each video and documented interactions with the diagram, including gestures. After this, we parsed the responses further in order to categorize the participants' reasoning based on their diagrammatic stage, (diagrammatization, exploration, observation, or abandonment of the diagram) and the type of representation (algebraic or geometric). This analysis allowed us to capture both aspects of our research purpose: relationships between algebraic and geometric reasoning, and the integration of diagrams and gesture into their geometric reasoning. Framed by this dual purpose, the results presented in this paper are based on Kelly and Abby's responses to four tasks: (a) Give a geometric explanation as to why $z\bar{z}=|z|^2$ is true. (b) Give a geometric explanation for the statement: If $\frac{1}{z}$ is defined by $\left(\frac{1}{z}\right)z=1$, why does it follow that $\frac{1}{z}=\frac{1}{re^{i\theta}}=\frac{1}{r}e^{-i\theta}$? (c) Give a geometric explanation as to why $\frac{\mathrm{Re}^{i\phi}}{re^{i\theta}}=\frac{R}{r}e^{i(\phi-\theta)}$ is true, and (d) Give a geometric explanation as to why $\frac{1}{x+iy}=\frac{x}{x^2+y^2}-i\frac{y}{x^2+y^2}$. In our results we provide details for tasks (a) and (c) but will present results of all four tasks. Task (b) was influential throughout the interview, possibly due to the amount of time devoted to it. In our results section the verbiage in bold represents utterances accompanied with gestures, which are described in parenthesis.

Results

In task (a), the girls correctly justify the task algebraically in an attempt to inform the geometric setting, but cannot completely justify the geometric aspect of the task. This was surprising because the participants repeatedly explained complex number multiplication as a rotation and dilation. Consider the following exchange, which occurred shortly after Abby drew z and an appropriate corresponding $\bar{z}$ on the diagram:

> *Abby: They have the same real number because the only thing that changes is the imaginary part.*
>
> *Kelly: Well, when you multiply the imaginary part goes away. Because it's the* ***middle terms*** **(**holds her left hand with fingers up as though holding a ball and then brings her fingers down and together**)**.

Abby: Mhm. And then it'll take you to the negative. Where the i squared…But I mean, because when you multiply them together it's a ***rotation*** (makes a rotation gesture again by moving her right index finger counterclockwise in an upper semicircle) *dilation.*

This conversation and their gestures seem to confirm that Abby and Kelly understand intuitively that $\bar{z}$ acts on z and how it does so, algebraically and geometrically. In the first exchange, the participants focus on the algebraic aspects, with Abby referencing the consistent real part and changing the imaginary part between z and $\bar{z}$. Abby's description speaks to some understanding of the algebraic relationship between z and $\bar{z}$, while Kelly's contribution suggests that she has a natural understanding of how complex multiplication works, at least in the algebraic setting. The final line in the above segment suggests that Abby recognizes the kinds of actions necessary to multiply two complex numbers geometrically, calling the two transformations collectively a "rotation dilation." Abby also produces a "rotation" gesture as shown in figure 1 as she utters this phrase. At this point in the task, Abby attempts to work out the algebraic details:

Abby: But it's a times a and a's the same for both of them. So that's why a squared is. And b squared. But one's negative and one's positive (gestures left for negative and right for positive). Because one is the conjugate of the other
Kelly: Don't know
Abby: I've got to do algebra

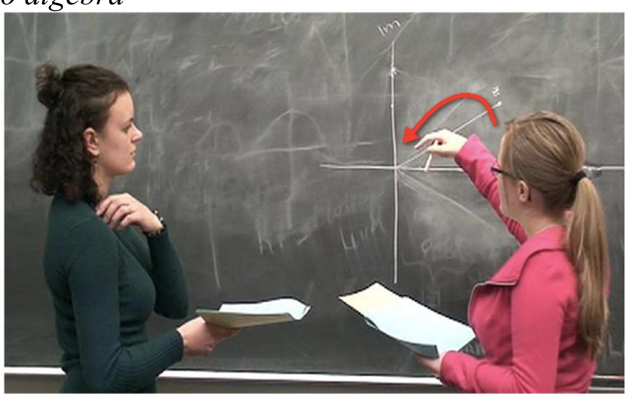


Fig. 1. Abby's rotation.

Initially Abby seems to make progress, uttering about little *a*'s and *b*'s, which most likely refers to the real and imaginary parts of z. Her gesturing seems to communicate that the real parts are different, although her speech suggests she believes it is the imaginary parts that differ. In this exchange and the next, Abby and Kelly effectively ignore their diagram. Because of their inability to make progress with geometry, Abby and Kelly both begin to reason algebraically. Kelly provides an algebraic proof and then attempts geometric reasoning. While neither Kelly nor Abby could reason geometrically before their algebraic digression, looking at the diagram seems to allow Kelly to recognize that the product of z and $\bar{z}$ can be thought of as z acting on $\bar{z}$. Abby also correctly explains how the magnitude should change, but neither seems to understand how the angles should be modified. Abby appears to utilize the diagram to ensure she understands Kelly's statements. While initially, Kelly agrees with Abby's belief that the vectors z and $\bar{z}$ form right angles, as soon as Kelly gestures toward the diagram, she realizes that the angle between z and $\bar{z}$ will not always be $90°$. Furthermore, Kelly articulates that the coefficient of *i* alone is what controls the measurement of this angle.

Finally, the interviewer prompts Kelly and Abby to describe in detail what exactly occurs when z is multiplied by $\bar{z}$. In their explanation, Kelly and Abby once again demonstrate that while they know that multiplication of two complex numbers should produce a "rotation dilation," neither appears to know exactly how to enact this operation in a particular context. Kelly again provides an appropriate explanation of how the magnitude changes and Abby fails to explain how multiplication rotates vectors. In the end, neither Abby nor Kelly articulates that the angles should be added together.

Abby and Kelly begin task (c) by noting that the given formula is algebraically correct, based on the rules of exponents. It was during this task that the pervasive influence of task (b) occurs for the first of many times. Contrasting it with task (b), Kelly realizes that task (c) concerns two arbitrary vectors.

Kelly: Yeah, doesn't matter now because we're not trying to get to anything.

Abby: Oh, yeah, that makes sense, sort of actually I don't have to do it, beginning. Well I was thinking we were going to have to actually do the act of multiplying the two like we did in the last problem. But that's not the case.

As they continue, Kelly revoices Abby's declaration that vector division is basically multiplication with fractions, and Abby pronounces that they will "have to draw it". In building their diagram, they marked φ and θ as the reference angles instead of the angles between the vectors and the positive real axis. After prompting, Kelly corrects the diagram, which allows her to reason through the task. Kelly makes good progress in determining the vector representation of the reciprocal of the denominator, paying particularly close attention to the angle of the reciprocal. While Kelly carefully locates the appropriate angle, she and Abby both appear to be lost regarding the appropriate magnitude for this new vector. This confusion may be partially due to task (b), which required finding the location of $\frac{1}{z}$ given an arbitrary vector z, and at this particular moment, finding a reciprocal is exactly what they are trying to do. Near the end of this exchange, Abby and Kelly discuss whether the vectors "have to go back to one" as in task (b). Kelly eventually realizes how to geometrically calculate the magnitude of their inverse vector. In particular, she realizes that a vector times its reciprocal will always be one – which they had explored in task (b).

After convincing Abby that the reciprocal is correct, both participants move on to "multiplying the two vectors" $\mathrm{Re}^{i\varphi}$ and $\frac{1}{re^{i\theta}}$. As Abby clarifies that they do have the correct angle for the "fraction" vector, she introduces another misconception regarding angles.

Abby: But it's not going to, it's going to go to negative one, because those are going to go to negative one.

*Kelly: No, because this would be a **negative x*** (taps near end of vector in 2nd quadrant)*, and a* ***negative** x* (taps near end of vector in 3rd quadrant).

Abby: Oh, I see what you're saying

Kelly: Like, on the graph it'll go to negative one, but numerically, it'll be a positive.

Kelly doesn't appear concerned by this potential problem, as she is confident that algebraically the product will result in positive one since the product of two negatives is a positive. As can be seen in the previous exchange, Kelly taps each of the vectors in turn as she speaks the word "negative" twice. As vectors themselves do not have a sign value, this pointing could suggest Kelly is attending in particular to the negative real parts of the indicated vectors.

Furthermore, prior to this exchange, she makes a “collapsing” gesture while explaining geometrically why she believes the product should result in negative 1 (see Figure 2).

Fig 2. Collapsing the product to -1.

Although Abby asked a geometric question and Kelly answered algebraically, Abby appears content with the explanation. The mismatch in answers between algebraic and geometric representations no longer seems to be a point of concern for either participant at this point. The interviewer, however, notices this discrepancy and asks for an explanation. In response, Kelly references task (b) again to justify her calculation of the appropriate reciprocal vector, this time providing explanations for finding both the new angle and the new magnitude. She then describes how to multiply the reciprocal vector by the vector in the “numerator” of the fraction, and finishes by describing multiplication in general as a rotation and a dilation. Abby agrees and gives an entirely algebraic explanation.

Discussion

Algebra seemed essential in helping our participants through the construction stage of diagrammatic reasoning as outlined by Bakker and Hoffman (2005), particularly for tasks (a), (c), and (d). For task (d) in particular the algebra seemed to elucidate which pieces were needed for the diagram and additionally initiated the change in forms. This may indicate that our participants, unlike Panaoura’s et al., (2006) or Danenhower’s (2006), did not have algebraic and geometric representations fully compartmentalized. Furthermore, our data suggest that once our participants understood the entities needed for their diagram, they were able to enact these pieces appropriately by treating some pieces as operators and others as operands. These enactments appeared to assist our participants to embody the geometric interpretations. For example, the embodiment of a “rotation dilation” did not come to fruition until they had opportunities to enact these transformations on their diagram using their fingers to represent the vectors and outstretched arms or shrinking movements to represent the dilation. By the end of the interview, illustrating the product of two complex numbers and the reciprocal of a complex number using diagrammatic reasoning and gesture became natural. We also noted that as we progressed further in the interview, dynamic gesture during the observation stage began to diminish, which could indicate that our participants were communicating mathematics they had embodied. Future studies may want to investigate if such gestures re-emerge as students tackle related but novel tasks. In teaching, it appears that requiring students to explicitly provide geometric interpretations resulted in a better understanding of the algebraic equations, their components, and the processes that allow one to justify the equations algebraically. Instructors could capitalize on this knowledge to create intentional teaching strategies, where they could not only model such reasoning but also highlight students' gestures, just as they might highlight students' work or verbal responses.

References

Bakker, A., & Hoffmann, H. G. (2005). Diagrammatic reasoning as the basis for developing concepts: A semiotic analysis of students' learning about statistical distribution. Educational Studies in Mathematics, 60, 333-358.

Châtelet, G. (2000). *Figuring space: philosophy, mathematics, and physics*. (R. Shore and M. Zagha, Trans.) Dordrecht, The Netherlands: Kluwer Academic Publishers.

Conference Board of the Mathematical Sciences. (2012). *The Mathematical Education of Teachers II Draft*. Retrieved from http://www.cbmsweb.org/MET2/MET2Draft.pdf.

Conner, M. E., Rasmussen, C., Zandieh, M., & Smith, M. (2007). Mathematical knowledge for teaching: The case of complex numbers. *Proceedings for the Tenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education.*

Council of Chief State School Officers & National Governors Association Center for Best Practices. (2010). *Common core state standards for mathematics.* Common Core State Standards Initiative. Retrieved from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Danenhower, P. (2000). Teaching and learning complex analysis at two-British Columbia universities. *Dissertation Abstract International 62(09).* (Publication Number 304667901). Retrieved March 5, 2011 from ProQuest Dissertations and Theses database.

Danenhower, P. (2006). Introductory complex analysis at two British Columbia universities: The first week – complex numbers. In F. Hitt, G. Harel, & A. Selden (Eds.) *Research in collegiate mathematics education VI*. Rhode Island: AMS.

de Freitas, E., & Sinclair, N. (2012). Diagram, gesture, agency: Theorizing embodiment in the mathematics classroom. *Educational Studies in Mathematics*, *80*(1-2), 133-152.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge: Belknap Press.

Panaoura, A., Elia, I., Gagatsis, A., & Giatlilis, G. P. (2006). Geometric and algebraic approaches in the concept of complex numbers. *International Journal of Mathematical Education in Science and Technology*, *37(6)*, 681-706.

Presmeg, N. C. (2006). Research on visualization in learning and teaching mathematics: Emergence from Psychology. In Á Gutiérrez, & P Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present, and future* (pp. 205-235). Rotterdam, Sense.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: author.

Needham, T. (1997). *Visual Complex Analysis*. Oxford: Oxford University Press.

Nemirovky, R., Rasmussen, C., Sweeney, G., & Wawro, M. (2012). When the classroom floor becomes the complex plane: Addition and multiplication as ways of bodily navigation. *Journal for the Learning Sciences*, *21(2)*, 287-323.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies of Mathematics, 22*, 1-36.

122

CURRENT AND FUTURE FACULTY MEMBERS' MATHEMATICAL KNOWLEDGE FOR TEACHING CALCULUS

Natasha Speer
University of Maine
speer@math.umaine.edu

Shahram Shawn Firouzian
University of Maine
Shahram_Firouzian@umit.maine.edu

Findings from research into "mathematical knowledge for teaching" have informed the design of preparation and professional development programs for K-12 teachers. At the college level there has been limited research into mathematical knowledge for teaching. We lack findings that demonstrate that expert teachers of college mathematics know and make use of knowledge beyond solely mathematical content.. The goal of this study is to examine the knowledge of student thinking possessed by mathematicians who teach calculus. Data come from interviews on student thinking about core calculus concepts.. Interviewees were research mathematicians who have been recognized for their teaching excellence and mathematics graduate students.. Findings demonstrate that the mathematicians were more able to identify known student difficulties as well as to describe common strategies students use to successfully solve the problems. Implications for research and professional development for novice college mathematics instructors are discussed.

Key words: teacher knowledge, knowledge of student thinking, mathematical knowledge for teaching, graduate students

Problem Statement

Concerns about enrollment and retention rates as well as the depth and breadth of calculus students' understanding sparked much activity over the past several decades (see, e.g., Bressoud, 2004, 2010; Lutzer, Rodi, Kirkman, & Maxwell, 2007; Lutzer & Maxwell, 2002). To meet this country's needs for scientists and engineers, we must find ways to increase the quality of students' understanding and the number of students who succeed in calculus.

There is now broader recognition at K-12 levels that in addition to knowledge of content, effective teaching relies on knowledge of (a) how students think and (b) mathematics that is "specialized" for the work of teaching (e.g., making sense of students' written or spoken work). Researchers have demonstrated that teachers with stronger knowledge of these sorts help students learn more mathematics content. These findings have prompted the K-12 education community to include these kinds of knowledge in professional development for teachers.

There is some evidence that these elements of "mathematical knowledge for teaching" also play roles in the teaching practices of college mathematics instructors, especially those practices needed for inquiry-oriented approaches to instruction (Speer & Wagner, 2009; Wagner, Speer, & Rossa, 2007). What the community lacks, however, is strong evidence that effective teachers of college mathematics possess this knowledge and use it in their instructional practices. Armed with information about the knowledge used by such instructors, professional development for novice college mathematics instructors (e.g., graduate students) could be designed to focus specifically on the development of such knowledge. This in turn could create better learning opportunities for students and lead to better achievement and retention.

The current study is focused on two research questions: What knowledge of student thinking and specialized content knowledge do experienced teachers of calculus possess? And, how does this compare to knowledge of novice teachers of calculus?

Research on Mathematical Knowledge for Teaching

Theoretical perspective

This project lies at the intersection of research on teachers' knowledge and on teachers' practices and was conducted using a cognitive approach. This approach has been used productively to examine teachers' knowledge and its roles in teaching practices (Borko & Putnam, 1996; Calderhead, 1991, 1996; Escudero & Sanchez, 2007; Schoenfeld, 2000; Sherin, 2002). In such an approach, knowledge is seen as a key factor influencing teachers' goals and the ways they work to accomplish those goals as they plan for, reflect on, and enact instruction.

Knowledge for teaching

No one questions the essential role that content knowledge plays in teachers' practices. However, such knowledge in itself is not strongly linked to student achievement (Ball, Lubienski, & Mewborn, 2001; Wilson, Floden, & Ferrini-Mundy, 2002). Research findings suggest that other types of knowledge play substantial roles in teachers' practices and learning opportunities they create for students.

Pedagogical content knowledge (PCK) refers to (among other content-specific things) knowledge of topics which typically cause students difficulty, the nature of those difficulties, and particularly useful examples for teaching (Shulman, 1986). Teachers' knowledge of the different strategies their students would use to approach problems is positively correlated with students' achievements (see, e.g., Fennema et al., 1996). For this project, analyses concentrate on knowledge of students' ideas (KSI), a subset of PCK used in the research noted above.

Specialized content knowledge (SCK) is a form of knowledge, not necessarily developed in ordinary mathematics courses, that enables teachers to engage in teaching tasks (Ball & Bass, 2000; Hill et al., 2005, 2004). SCK is used to follow students' thinking, evaluate validity of student-generated strategies, and make sense of student-generated solution paths (Hill, Ball, & Schilling, 2008). Teachers' SCK has been shown to be positively related to student achievement gains in elementary mathematics (e.g., Hill et al., 2005).

Research Design

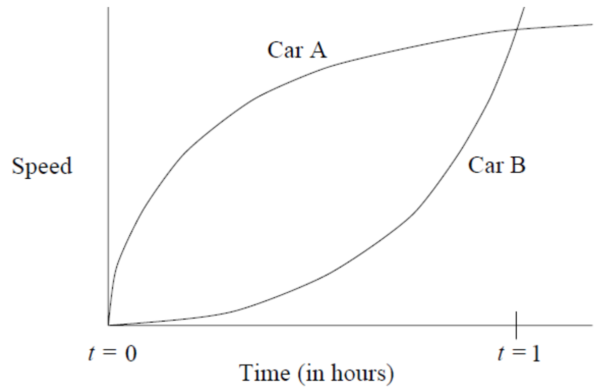
Data come from task-based individual interviews with research mathematicians who have been recognized for their excellence in teaching (e.g., nominated for or won a teaching award) and graduate students in mathematics with less than two years of calculus teaching experience.

Tasks were taken from or modeled after tasks used in research on student thinking about limit, function (as it appears in calculus), derivative, and integral. Interviews consisted of three parts per task for each interviewee: Solve the task and describe the solution; Describe how students would solve the task, including difficulties/mistakes they might make and correct/incorrect ways of thinking they might display.

Data analysis was guided by research on student thinking but also relied on methods from Grounded Theory (Strauss & Corbin, 1990). Findings from research on student thinking were used to characterize the extent to which participants were knowledgeable of student thinking. Methods from Grounded Theory were then used to identify themes and to detect differences between findings from mathematicians and novice instructors.

Here we present findings from one task. Borrowed from Carlson (1998), this question taps into students' abilities to interpret graphical information about two functions (see Figure 1).

The given graph represents speed vs. time for two cars. (Assume the cars start from the same position and are travelling in the same direction.)



Question: State the relationship between the position of car A and car B at t = 1 hour. Provide an explanation for your answer.

Figure 1. Interview Task related to students' difficulties with function graphs

Findings

All participants generated correct solutions to the task. All participants were also able to describe at least some correct ways of thinking that students might use. For example, one participant said students might get the solution by thinking "speed is greater always … or if the velocity is greater always … the displacement at the end is going to be greater." All participants were also able to describe some incorrect ways students might think about the task. These ways included ones documented in research on student thinking using this task. For example, participants noted that students might interpret the graph as if the dependent variable was distance instead of speed and conclude that the two cars have traveled the same distance at t = 1. In describing this kind of thinking, one participant stated, "A lot of students probably would say … the two graphs are intersecting at time equals to 1 so they are equal…and forget about what it is that is equal."

There were differences in the extent to which the two populations (research mathematicians and graduate students) were able to generate possible student ways of thinking about the task. The mathematicians were generally able to describe more distinct ways than the graduate students. The mathematicians appeared to possess more knowledge of student thinking (PCK) from their experiences working with students. They also appeared more able to hypothesize other possible ways of thinking based on their knowledge of mathematics (e.g., using their SCK to create hypothetical ways one might approach the tasks even if they had not actually seen such an approach). Graduate students, in contrast, had a narrower set of ways of thinking from which to work and were less successful in generating a variety of potential approaches, sometimes focusing just on small variations to one approach they knew. For example, one graduate student said, "they might not, [they] can't read … the graph they don't know what does this graph represent [they] just don't know how to read this graph I think.." He tried, with limited success, to construct possible students' mistakes based on an inability to read the graph but was not able to describe other kinds of difficulties students might have.

Conclusions

The graduate students had some knowledge of students' difficulties and students' strategies to solve the above mentioned problem. However, when asked to construct possible students' responses, graduate students had more difficulty than the mathematicians. All participants had the necessary subject matter knowledge to solve the task, however the mathematicians had a robust knowledge of students' ideas, their difficulties with certain concepts, and were able to generate new ways students might think by using their content knowledge and specialized content knowledge.

Implications for Research and Practice

By comparing the knowledge of experienced college mathematics instructors to that of novices (and to what is known from research on student thinking for particular topics), we can identify areas where novice instructors might profit from professional development. The experienced instructors said they had learned about student thinking from their interactions with students and from examining students' written work. Armed with findings about what novices do and do not know, professional development can be designed to help graduate students learn as much as possible from their interactions with students and student work so they can begin their faculty careers equipped with as much knowledge of student thinking as possible. This in turn may enable them to more quickly develop into accomplished teachers.

It also appears that the experienced instructors were well versed in the practices of anticipating or interpreting student ways of thinking by drawing on their knowledge of the mathematics content and their knowledge of student thinking. Further analysis may shed light on how experts do this particular kind of teaching-related work and may contribute to theories of how teacher generate new knowledge for teaching while engaged in the work of teaching.

Discussion Questions

1. In addition to examining samples of student work, are their other activities we could do in the interviews that would reasonably simulate the authentic work of teaching and generate data on mathematical knowledge for teaching?
2. What reasonably compact approaches might be best for presenting both the breadth and depth of an individual participant's knowledge? What approaches might be best for describing these things for the two populations?

References

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics*. Westport, CT: Ablex.

Ball, D. L., Lubienski, S., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of Research on Teaching* (pp. 433–456). Washington, DC: American Educational Research Association.

Borko, H., & Putnam, R. T. (1996). Learning to Teach. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 673–708). New York: Macmillan Library Reference USA: Simon & Schuster Macmillan.

Bressoud, D. (2004). The changing face of calculus: First-semester calculus as a high school course. *FOCUS*, *24*(6), 6–8.

Bressoud, D. (2010). The rocky transition from high-school calculus. *The Chronicle of Higher Education*.

Calderhead, J. (1991). The Nature and Growth of Knowledge in Student Teachers. *Teaching and Teacher Education*, *7*(5/6), 531–535.
Calderhead, J. (1996). Teachers: Beliefs and Knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 709–725). New York: Macmillan Library Reference USA: Simon & Schuster Macmillan.
Carlson, M. (1998). A Cross-Sectional Investigation of the Development of the Function Concept. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in Collegiate Mathematics Education* (Vol. 1, pp. 114–162). Washington, DC: American Mathematical Society.
Escudero, I., & Sanchez, V. (2007). How do domains of knowledge integrate into mathematics teachers' practice? *The Journal of Mathematical Behavior*, *26*(4), 312–327. doi:10.1016/j.jmathb.2007.11.002
Fennema, E., Carpenter, T., Franke, M., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics Instruction. *Journal for Research in Mathematics Education*, *27*(4), 403–434.
Hill, H., Ball, D. L., & Schilling, S. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education*, *39*(4), 372–400.
Hill, H., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371–406.
Hill, H., Schilling, S., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, *105*(1), 11–30.
Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 CBMS Survey. Washington, DC: American Mathematical Society.
Lutzer, D., & Maxwell, S. (Eds. . (2002). *Statistical Abstract of Undergraduate Programs in the Mathematical Sciences in the United States: Fall CBMS Survey*. Providence, RI: American Mathematical Society.
Schoenfeld, A. H. (2000). Models of the Teaching Process. *Journal of Mathematical Behavior*, *18*(3), 243–261.
Sherin, M. (2002). When teaching becomes learning. *Cognition and Instruction*, *20*(2), 119–150.
Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.
Speer, N., & Wagner, J. (2009). Knowledge Needed by a Teacher to Provide Analytic Scaffolding During Undergraduate Mathematics Classroom Discussions. *Journal for Research in Mathematics Education*, *40*(5), 530–562.
Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park: Sage.
Treisman, U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *College Mathematics Journal*, *23*(5), 362–372.
Wagner, J., Speer, N., & Rossa, B. (2007). Beyond mathematical content knowledge: A mathematician's knowledge needed for teaching an inquiry-oriented differential equations course. *Journal of Mathematical Behavior*, *26*, 247–266.
Wilson, S., Floden, R., & Ferrini-Mundy, J. (2002). Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education*, *53*(3), 190–204.

123

TECHNOLOGY AND ALGEBRA IN SECONDARY MATHEMATICS TEACHER PREPARATION PROGRAMS

Eryn M. Stehr and Lynette D. Guzman
Michigan State University

Most recently, the Conference Board of the Mathematical Sciences has advocated for incorporating technology in secondary mathematics classrooms. Colleges and universities across the United States are incorporating technology to varying degrees into their mathematics teacher preparation programs. This study examines preservice secondary mathematics teachers' opportunities to expand their knowledge of algebra through the use of technology and to learn how to incorporate technology when teaching algebra in mathematics classrooms. We explore the research question: What opportunities do secondary mathematics teacher preparation programs provide for PSTs to encounter technologies in learning algebra and learning to teach algebra? We examine data collected from a pilot study of three Midwestern teacher education programs conducted by the Preparing to Teach Algebra (PTA) project investigating algebra. Our data suggest that not all secondary mathematics teacher preparation programs integrate experiences with technology across mathematics courses, and that mathematics courses may provide few experiences with technology to PSTs beyond strictly computational.

Key words: Algebra and Algebraic Thinking, Technology, Preservice Teacher Education, High School Education

This study explores opportunities provided by secondary mathematics teacher preparation programs for preservice teachers (PSTs) to expand their knowledge of algebra through the use of technology and to learn how to incorporate technology when they teach algebra. We explore the following research question: *What opportunities do secondary mathematics teacher preparation programs provide for PSTs to encounter technologies in learning algebra and learning to teach algebra?* These opportunities might include using or learning about a variety of algebra-appropriate technologies, as well as thinking critically about technology use. In this study, we define technology narrowly as electronic tools and software. This study will not focus on physical tools such as manipulatives, chalkboards, or dry erase boards, although we acknowledge that these tools are also important technologies that can be useful for teaching and learning mathematics.

Context

Technology use in K-12 education has become practically universal in the past few decades. Many scholars suggest that use of technological tools in the classroom could contribute to reducing inequities in education. For example, Pomerantz (1997) argued: "....Calculators serve as an equalizer in mathematics education" (p. 5). Technology use, however, has led to a so-called digital divide (Reich, Murnane, & Willett, 2012). Attewell and Gates (2001) described the digital divide as two-fold: a division of access and of use. Federal funding has mitigated issues of access; however, there is a growing recognition of disparity in technology use in schools (Attewell & Gates, 2001). Thus, a focus shifts from supplying schools with technology to the highly effective ways in which technology can be (but is not usually) used.

Both secondary mathematics content standards and teacher preparation standards have emphasized the importance of developing PSTs' abilities to choose and use educational technologies. Standards developed for teacher preparation program accreditation agencies, such as National Council for Accreditation of Teacher Education (NCATE: NCTM, 2012) and Interstate Teacher Assessment and Support Consortium (InTASC: CCSSO, 1995), recommended that PSTs develop the abilities to critically evaluate and strategically use technology. In addition, the Conference Board of the Mathematical Sciences (CBMS) emphasized the importance of PSTs' preparation to use technology in *Mathematics Education of Teachers II* (CBMS, 2012).

Algebra plays a prominent role in mathematics education reform efforts because it is valued as an important subject in mathematics. In terms of equity issues related to mathematics education, algebra has long been seen as a gatekeeper for post-secondary opportunities (e.g., Moses, Kamii, Swap, & Howard, 1989). Particularly in the United States, preparing future secondary mathematics teachers to teach algebra has gained importance as more states include algebra as a high school graduation requirement (Teuscher, Dingman, Nevels, & Reys, 2008). Consideration of state education websites verifies that at least 38 states currently include mathematics courses with algebra as a necessary high school graduation requirement. Algebra is also being offered earlier in some states. In 1990, only 16% of all eighth-graders were enrolled in algebra, and this increased to 31% by 2007 (Loveless, 2008). The emphasis of algebra in mathematics education, along with increasing use of technology in the classroom, highlights the need to support future mathematics teachers in learning algebra with technology and learning to teach algebra with technology.

To use technology effectively to support the teaching of algebra, CBMS (2012) argued that experience with technology "should be integrated across the entire spectrum of undergraduate mathematics" (pp. 56-57) and PSTs should have opportunities to see teaching with technology modeled in their own mathematics coursework (CBMS, 2012). PSTs need to become familiar with a variety of technological tools used in a variety of ways, including computational tools, problem-solving tools, and tools for exploring mathematical ideas (CBMS, 2012; NCATE, 2012; InTASC, 1995).

Method

This study is part of a larger mixed-methods study, *Preparing to Teach Algebra (PTA),* that explores opportunities provided by secondary mathematics teacher preparation programs to learn algebra and to learn to teach algebra. The *PTA* project consists of a national survey of secondary mathematics teacher preparation programs and case studies of five universities. The current study is a qualitative analysis of opportunities provided to PSTs to encounter technology in learning algebra and learning to teach algebra based on data gathered during the pilot study of the *PTA* project.

In the pilot study, the *PTA* project chose three secondary mathematics teacher preparation programs for convenience. University A is a medium-size university with Carnegie classification of RU/H (Research University with high research activity). Universities B and C are large universities, both with Carnegie classification RU/VH (Research University with very high research activity). The programs at Universities A and C are four-year programs, and the program at University B is a five-year program.

We compiled data by conducting one focus group and five instructor interviews at each site and collected corresponding instructional materials from each instructor we interviewed. Of the five instructor interviews at each site, we included three mathematics courses and two mathematics education courses. We selected courses based on potential for algebra content, availability and course type. Among other questions, we asked instructors which types of technologies they used in a particular course; we also analyzed their course materials. Three or four students who had completed, or had almost completed, their student teaching requirement participated in each focus group. We asked PSTs about their required or shared experiences. They confirmed a list of program requirements and identified required courses that incorporated technology in learning algebra or learning to teach algebra.

Because this study uses pilot data from a larger study, one limitation is that instructor interviews were restricted to five courses at each site that were not representative of the entire teacher preparation program. Additionally, we chose courses based on their likelihood to contain algebraic content and not for a focus on technology. As a result, we may have missed some courses that focus on technology in secondary mathematics. To balance this limitation, we used information from each focus group and course descriptions obtained from each school website to create an outline sketch of technology use for each program.

In our sketch of technology use, we first identified examples of technology use in algebra from instructor interviews, focus group interviews, or from the instructional materials. We analyzed each example according to five characteristics of experiences: activity type, types of technology use, algebraic topics, type of technology, and whether PSTs have the opportunity to think critically about choice and use of technology.

Results

Due to space limitations, we give a brief overview of what we have learned in this proposal and more detailed results will be provided in the presentation. Across all universities, we found a total of 28 examples of algebraic topics using technology, with 8 found in mathematics content courses and 20 found in mathematics education courses. This count excludes numerous examples in a Differential Equations course at University C, which involved a computer lab component. A descriptive list of algebraic topics and in which courses examples were found (M for Mathematics Courses and ME for Mathematics Education Courses) are shown in Table 1.

Table 1. *Algebra topics using technology identified per university and by mathematics or mathematics education courses.*

Algebraic Topics	University A	University B	University C
Generalizing Patterns		ME	
Maximum Area Problem		ME	
Ratios and Proportion		ME	ME
Modeling with Equations			ME
Functions and Multiple Representations	ME	ME	M
Linear Functions (e.g., families, slopes)		ME	ME
Systems of Linear Equations	M, ME		
Parametric Equations		ME	
Logarithmic Functions	ME		
Matrices	M		
Topics from Calculus	M	M	
Topics from Differential Equations			M
Modular Arithmetic	M		
Extensions on Rational Numbers	M		

University A requires twelve mathematics courses and four mathematics education courses. The Mathematics Department offers the mathematics education courses. Teaching Mathematics with Technology is one of the required mathematics education courses. One instructor said, “We think [technology is] crucial.” The number of mathematics courses that include technology use in their course descriptions (six of twelve) and the one mathematics education course focused on teaching with technology supports this statement of belief. Overall, student and instructor responses indicated technology was used in mathematics courses primarily as a computational tool, while mathematics education courses supported a greater variety of types of uses of technology, including some critical evaluation of technology for algebra teaching. One example of technology use for algebra comes from an activity in the Technology in Secondary Mathematics course in which PSTs “investigate graphing utilities and think about what are the features of graphing utilities that would … make one more desirable than another.” In this assignment, neither instructors nor PSTs necessarily used technology, but students thought critically about uses and types of graphing utilities. A second example came from the Structure of Algebra course instructor, saying “when we talk about cryptography I'll bring in Mathematica... if you want to do RSA cryptography in any sort of realistic way, you want to use... you know, RSA relies on a number that's a product of large primes. …So you're doing … arithmetic mod some huge number.”

University B requires eight mathematics courses and four mathematics education courses. The College of Education offers the mathematics education courses. PSTs in the focus group marked some use of computer software in Calculus III, geometry courses, and statistics courses, as well as multiple technologies in the four mathematics education courses. One student in the focus group stated, “…tools for me is by far the biggest weakness… even when we did use them it was pretty rare.” Instructors of Linear Algebra, Analysis, and the Capstone course stated that they did not use technology in class, except rarely to check a calculation. The Linear Algebra instructor explained, “I don't think it is a good idea to use calculator or computer software… you want them to do it by hand.” The mathematics education courses used multiple instructional and mathematical technologies to support algebra topics. Specific mathematical technologies included GeoGebra, spreadsheets, graphing calculators, and the occasional use of Geometer’s Sketchpad. One instructor explained that he chose to use technology because “…[the PSTs] see things mathematically they didn't see before and it helps them see the value of engaging in those sorts of tasks with their students….” Overall, student and instructor responses indicated that few mathematics courses used technology, while mathematics education courses integrated a variety of technologies to support PSTs teaching and learning of mathematics as well as PSTs critical evaluation of technologies. One example of technology use in algebra was in the 2nd Secondary Methods course, the instructor introduced students to the “Ships in the Fog” task (based on the crash of the Stockholm and Andrea Daria) through a newsreel video of the wreck, solving the problem three ways (the worksheet calls for graphing calculator use), discussing the task on the Wiki, and then reading the “Ships in the Fog” case. The technological tools would be the video clip, graphing calculator, and Wiki site. The video clip connects the task to the real world situation, the role of the graphing calculator (computation only or also exploration?) is not clear, and the Wiki site provides a venue for reporting their own solutions and discussion other solution strategies.

University C requires twelve mathematics courses and two mathematics education courses. The College of Education offers the mathematics education courses. An Educational Technology course is a required course in the program but does not focus on mathematics. The mathematics department at University C does not allow graphing calculators on exams. PSTs indicated four mathematics courses in which they used computer software or clickers and did not indicate technology use in mathematics education courses. PSTs stated that they did not learn to use certain technologies despite needing them later in field instruction. For example, one PST wrote along the list of courses, “no graphing or non-graphing calculator allowed.” The Abstract Algebra and Differential Equations instructors indicated rare use of technology in lectures; however, the Differential Equations course included a computer lab component using MatLab. Although the mathematics education course instructors did not emphasize technology in their instruction, one assignment did require students to revise a previously written lesson plan to “include technologies that *enhance the teaching and learning of mathematics*,” and to discuss their rationale for inclusion. Overall, student and instructor responses indicated several mathematics courses used technology, while mathematics education courses supported critical evaluation through choice and justification of technology for mathematics teaching, although the courses themselves did not integrate technology use. One example of technology use in algebra was one activity in Differential Equations called “The Swaying Building.” This activity had the goal: “Determine a model of the swaying of a skyscraper; estimating parameters.” In this activity, PSTs build a mathematical model and use representations to investigate mathematical ideas, by comparing tabular and graphical representations (in Figure 5 below) of “measurements of displacement as a function of time of the building reacting to two different shocks.”

Discussion

Contrary to CBMS (2012) recommendations, our data suggest that not all secondary mathematics teacher preparation programs integrate experiences with technology across mathematics courses. We found that mathematics education courses integrate technology into instruction and learning more commonly than mathematics courses. Even in mathematics courses that use technology, our data suggest that PSTs have fewer opportunities to see and use a variety of technological tools and that PSTs are more likely to see or use technologies only as computational tools. With respect to specific experiences using technology in learning and learning to teach algebraic topics, according to our data, mathematics education courses provide the bulk of these experiences.

We heard concerns from both mathematics and mathematics education instructors that technology would impede PSTs’ learning. Some mathematics education instructors argued, to the contrary, that use of technology enabled PSTs to increase their understanding of algebra topics in ways that were not possible otherwise. One explanation of this difference in instructors’ viewpoints might lie in whether instructors used technology only as a practical expedient.

Further research should be done to investigate ways technology can be used more effectively in algebra to support future teachers’ understanding of algebra as well as their abilities to use technology more effectively in their own classrooms.

Endnote

This study comes from the *Preparing to Teach Algebra* project, a collaborative project between groups at Michigan State (PI: Sharon Senk) and Purdue (co-PIs: Yukiko Maeda and Jill Newton) Universities. This research is supported by the National Science Foundation grant DRL-1109256.

References

Attewell, P., & Gates, L. (2001). Comment: The First and Second Digital Divides. *Sociology of Education*, *74*(3), 252–259.

Conference Board of the Mathematical Sciences [CBMS]. (2012). *The Mathematical Education of Teachers II*. Providence RI and Washington DC: American Mathematical Society and Mathematical Association of America.

Council of Chief State School Officers [CCSSO]. (2011). *Interstate Teacher Assessment and Support Consortium (InTASC) Model Core Teaching Standards: A Resource for State Dialogue*. Washington, DC: Author.

Loveless, T. (2008). *The misplaced math student: lost in eight-grade algebra*. The Brown Center Report on American Education.

Moses, R. P., Kamii, M., Swap, S. M., Howard, J. (1989). The Algebra Project: Organizing in the spirit of Ella. *Harvard Educational Review*, 59(4), 423-443.

National Council of Teachers of Mathematics (NCTM). (2012). *NCATE/NCTM Program standards: Programs for initial preparation of mathematics teachers.* Retrieved October 12, 2012, from http://www.nctm.org/uploadedFiles/Math_Standards/NCTM%20NCATE%20Standards%20Revision%20and%20Math%20Content%20-%20Secondary.pdf

Pomerantz, H. (1997). *The role of calculators in math education (Research compiled under the direction of Bert Waits)*. Columbus: The Ohio State University.

Reich, J., Murnane, R., & Willett, J. (2012). The State of Wiki Usage in U.S. K-12 Schools: Leveraging Web 2.0 Data Warehouses to Assess Quality and Equity in Online Learning Environments. *Educational Researcher*, *41*(1), 7–15. doi:10.3102/0013189X11427083

Teuscher, D., Dingman, S. W., Nevels, N. N., & Reys, B. J. (2008). Curriculum Standards, Course Requirements, and Mandated Assessments for High School Mathematics: A Status Report of State Policies. *NCSM Journal*, 49.

PERCEPTIONS IN ABSTRACT ALGEBRA: IDENTIFYING MAJOR CONCEPTS AND CONCEPT CONNECTIONS WITHIN ABSTRACT ALGEBRA

Ashley L. Suominen
University of Georgia

Abstract algebra is recognized as a highly problematic course for most undergraduate students. Despite these difficulties, most mathematicians and mathematics educators affirm its importance to undergraduate mathematical learning. The goal of this research was to establish a list of the important concepts in abstract algebra as perceived by graduate students in mathematics and understand how they believe these concepts are related. Through an interview study, the students' perceptions of abstract algebra were analyzed through the development of concept maps. Participants' concept images and concept definitions are highlighted in this article to understand their concept perceptions. The results revealed graduate students had great difficulty articulating what they learned and their concept definitions. Consequently, they had differing views of major concepts and relationships within the course. Their concept images regarding perceived concept importance seemed to be equated to the amount of time their class spent discussing that concept.

Key words: Abstract algebra, Concept maps, Connections, Concept image, Concept definition

Introduction

It is widely acknowledged that abstract algebra is an essential part of undergraduate mathematical learning (e.g., Gallian, 1990; Hazzan, 1999; Selden and Selden, 1987), and yet it is also known for its high level of difficulty at the collegiate level. Many undergraduate and graduate students, including prospective teachers, struggle to grasp even the most fundamental concepts of this course (Dubinsky et al, 1994). For most of these students, abstract algebra is the first time they experience mathematical abstraction and formal proof. It is often the first course in which teachers expect students to "go beyond learning 'imitative behavior patterns' for mimicking the solution of a large number of variations on a small number of themes (problems)" (Dubinsky et al., 1994, p. 268) by requiring proofs to explain abstract theories and ideas. In particular, students are expected to mentally construct new objects based on a list of properties and then operate on these objects. However, simply being exposed to these abstract concepts does not imply the development of mathematical meaning. Von Glasersfeld (1991) affirmed this predicament: "Reflecting upon experiences is clearly not the same as having an experience" (p. 2). Students must take an active role in the learning process by building on their past mathematical knowledge to make sense of abstract concepts.

In his dissertation, Cook (2012) asserted the difficulty students experience in abstract algebra is due to the lack of established connections between undergraduate mathematics and school mathematics. He affirmed that prospective teachers "do not build upon their elementary understandings of algebra, leaving them unable to communicate traces of any deep and unifying ideas that govern the subject" (p. xvi). These conjectures imply that undergraduate professors must be able to not only convey an abstract idea to students but also provide students the opportunity to build mathematical meaning upon these abstractions. Fennema and Franke (1992) supported this theory: "If teachers do not know how to translate those abstractions into a form that enables learners to relate the mathematics to what they already know, they will not learn with understanding" (p. 153). Thus, we can only expect undergraduate students to really access the benefits of this course through complete comprehension by connecting abstract theory to past knowledge and ideas to aid in the construction of mathematical meaning. The purpose of this research was two-fold: 1) formulate a list of important concepts for the collegiate course abstract algebra, and 2) recognize students' perceived relationships or connections existing within these concepts. More specifically, three graduate students in mathematics were interviewed to discuss their perceived list of important concepts in abstract algebra as well as relationships between topics within the course.

Literature Review

Despite an increasing amount of research on teaching and learning collegiate mathematics, few studies concentrate solely on abstract algebra. The existing abstract algebra research concentrates on student learning (e.g. Asiala, Brown, Kleiman, & Mathews, 1998; Brown, DeVries, Dubinsky, & Thomas, 1997; Leron, Hazzan, & Zazkis, 1995) and teaching methods (e.g. Asiala, Dubinsky, Mathews, Morics, & Oktaç, 1997; Freedman, 1983; Pedersen, 1972). The former research primarily utilized the APOS (Action-Process-Object-Schema) theoretical framework to analyze student constructions of course topics. For example, Brown, DeVries, Dubinsky, and Thomas (1997) used APOS theory to examine how abstract algebra students understood binary operations, groups, and subgroups. In this study the researchers implemented an ACE teaching cycle (Activities, Class discussion, and Exercises) with computer activities using the language ISETL in order to aid in students' construction of aforementioned topics. They concluded that their pedagogical approach seemed reasonably effective. Similarly, the impact of pedagogical practices influenced by a constructivist perspective was investigated. For instance, Freedman (1983) introduced a unique lecture-based method that progressively required undergraduate abstract algebra students to take an active part in their learning through teaching. In this three-stage teaching method students initially learn through traditional lecture, then in the second stage students are required to complete a project as well as do a little of the teaching, and finally students are solely responsibly to design objectives and prepare all lectures. Through this active participation in the teaching process, students were able to gain a strong understanding of the topics.

Much of the current abstract algebra research affirms students' difficulties in learning fundamental concepts in group theory (e.g. Asiala, Dubinsky, Mathews, Morics, and Oktac, 1997; Larsen, 2004, 2009; Leron, Hazzan, and Zazkis, 1995) with little attention to rings or fields (Cook, 2012). Likewise, students' difficulties in proof writing have also been highlighted in the context of group theory (e.g. Hart, 1994; Selden and Selden, 1987; Weber, 2001). To date, there has been very little research on what concepts are deemed important in abstract algebra and why. Moreover, students' perceptions of concept importance and concept connections have not thoroughly been studied.

Theoretical Perspective

At the basis of this exploratory research study is the belief that students construct their own understanding of mathematics. A constructivist theoretical framework was thus utilized in suggesting that students construct their own reality of the mathematics they are learning (von Glasersfeld, 1989). More specifically, students develop concepts through a series of mental processes that largely depend on past experiences. The development of understanding of a concept often includes two components: concept image and concept definition. Tall and Vinner (1981) defined concept image as "the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes." For instance, students learning abstract algebra may construct mental pictures of specific algebraic structures in attempting to understand them. These pictures may include a list of properties, memories of class activities, or relations to past related concepts. The commonly accepted words used to describe these concept images are called a concept definition. Students, in particular, can define concepts either from rote memorization or through personal reconstruction of their concept images. Even though their personally constructed concept images seem sensible to the student, their understandings of the concepts may differ from the formally accepted definition. Thus, the concept images and concept definitions constructed by undergraduate and graduate students learning abstract algebra may or may not correspond to those taught in the course or found in the textbook.

Methodology

This research employed a semi-structured interview protocol with both open-ended questions and construction tasks (Patton, 2002; Taylor & Bogdan, 1984; Zazkis & Hazzan, 1999). Each interview was audio recorded and ran approximately 45-60 minutes in length in a private room to ensure confidentiality. After the interviews were complete the audio was transcribed within a week of the interview. Participants were chosen based on recently (within a year or less) being enrolled in the master's level abstract algebra course and being accepted into at least the master's level mathematics graduate program. While undergraduate students typically take abstract algebra, graduate students were chosen to provide an additional level of expertise. Three students (pseudonyms: Andrew, April, and Heather) participated in this research study. Each student had taken three lecture-based abstract algebra courses—an introductory course as an undergraduate and a yearlong sequence of two courses as a graduate student.

Since the purpose of this research study is to gain insight into graduate students' perspectives of abstract algebra, one of the central foci of the interview was the creation of concept maps. These maps allowed the participants and researcher to visually understand described relationships between concepts. Novak and Cañas (2008) and Trochim (1989) largely contributed to the overall research design of this activity. First, each participant was given index cards (or post-it notes) and asked to write any important or key concepts of abstract algebra on a card (one per card). When he or she was finished with this task, the participant was asked to explain each concept. Next, participants were asked to visually represent any conceptual relationships between these topics by placing their concept cards on a sheet of poster board and drawing lines or arrows between concepts that have some type of relationship. After each participant completed a concept map, he or she was asked to explain why each line was drawn.

Grounded theory was utilized when analyzing the data. Thus, the data was first collected, coded, grouped by concepts, categorized, and then the theoretical results were formulated (Charmaz, 2000). The transcribed interview responses were analyzed thematically (Charmaz, 2000; Patton, 2002; Taylor & Bogdan, 1984) focusing on the students' constructed knowledge of abstract algebra. The perceived significant concepts and connections among topics in the course were explicitly emphasized.

Results

As to be expected, each of the mathematics graduate students had a differing concept image and concept definition of major abstract algebra concepts. When asked to identify these concepts, April and Heather equated the time spent in class to the importance of the concept. April stated, "I think that fields are very important because we spent a lot of time discussing the different properties of fields and the different types of fields… So I felt it was really important."

Likewise, Heather repeatedly defined concept importance by how long the professor discussed it in class. Andrew, on the other hand, relied on his perceived usefulness of a certain concept to determine major concepts. When asked to describe ring theory Andrew stated:

> It's like you encounter rings first from like the first time you encounter math to be like the real numbers. We actually use them in our real life and everything, so in a way like this concept of rings kind of formalizes our understanding of what everything actually means.

However, despite the varying concept images associated with concept importance, there were five identified concepts that were mentioned by all three students: groups, rings, fields, Galois theory, and isometries with geometric applications. A complete summary of the perceived important concepts of each student is found in Figure 1.

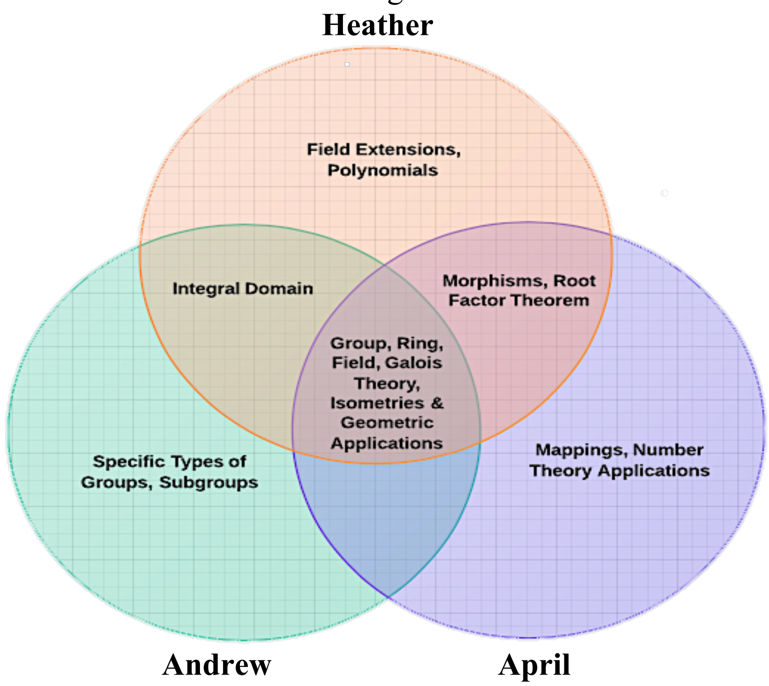
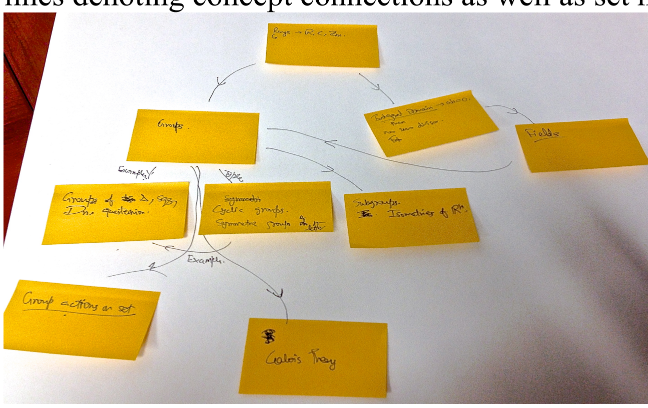


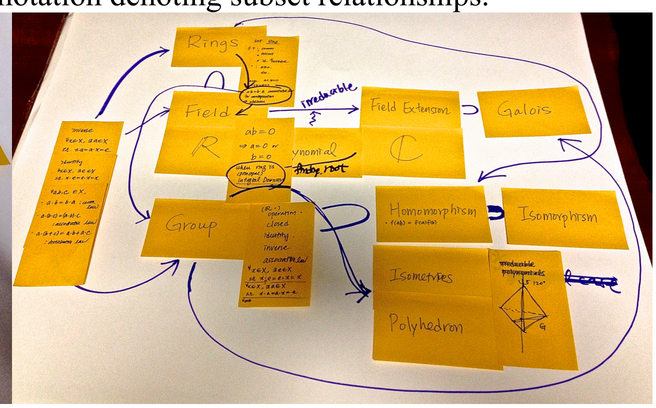
Figure 1: Identified important concepts

In general the mathematics graduate students had difficulty articulating their concept images about content learned in their abstract algebra courses. Despite all of the students acknowledging the intuitive nature of rings, none of them were able to articulate the complete formal definition of a ring. April's definition most closely aligned with the formal definition in classifying a ring as a set with two operations following seven axioms, but she did not articulate what those axioms were. Heather's concept image of a ring was similar to April's in that she viewed the algebraic structure in terms of axioms. However, Heather did not fully define a ring due to declared confusion between what those axioms were and how many existed for rings. On the other hand, Andrew's concept image of a ring was slightly different. His concept definition of a ring included the operations and properties, but also included the notion of a map. He stated, "It is something like you have a map, you have commutativity over addition, associativity over addition, and you have additive identity, you have multiplicative identity." When asked to describe a group or a field, the students seemed to have less robust concept images of these concepts than a ring. In fact, only April provided concept definitions for a group and field that closely aligned with the formal definitions. The other two graduate students had difficulty doing so. Heather responded, "Gosh. I think I am confused" and "The funny thing is I just totally, I just don't remember what a field was." In fact, this student repeatedly asked me during the interview to define group and field for her since she could not remember despite earning As in all of her courses. Andrew's concept image of a group related heavily on his definition of a ring, "From rings we can get groups. Kind of like subsets of rings are groups because we just have one operation." However, he also acknowledged, "A field is something I just can't get used to it." All participants provided a personal reconstruction concept definition of isometries and geometric applications, whereas no participant could accurately define Galois theory.

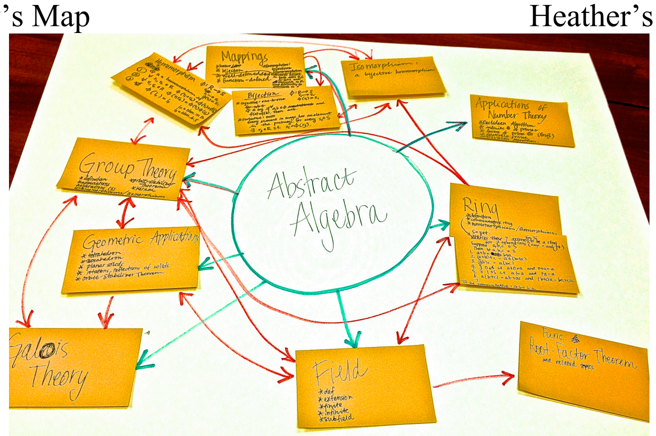
The concept images and concept definitions of perceived relationships between identified important concepts, as seen in the created concept maps, were quite diverse despite the fact that the participants took the same course (Figure 2). Andrew described his concept image of the connections between concepts as a "hierarchical structural" flow chart. When asked to describe the arrows drawn between concepts, he admitted to not fully grasping how concepts in abstract algebra were built upon each other, but he knew they were all somehow related. His description of a subset relationship between groups and rings aligned with this concept image as he stated, "Because rings are the more generated thing with two operations, addition and multiplication, so a group is kind of like throwing one of the operations out." Likewise, Heather's concept image of the connection between groups and rings also included a subset relationship, but she seemed unclear about how the subset relationship worked. She questioned whether a group is a subset of rings or rings a subset of groups. On the other hand, April's description of concept connections linked major concepts (rings, fields, groups, mappings, etc.) to the similarities and applications of her constructed concept definitions. Heather described her map as a web of concepts with lines denoting concept connections as well as set notation denoting subset relationships.



Andrew's Map



Heather's Map



April's Map

Figure 2: Graduate students' concept maps

Despite the students' differing concept images regarding the perceived concept connections, two of the graduate students constructed concept images of these connections by the order the concepts were discussed in class. This emphasis on classroom activities parallels how the students defined concepts importance. April explained when asked to elaborate on her concept maps, "The reason I have a bidirectional arrow between fields and rings was because we discussed fields after rings." Likewise, Heather described her arrow between ring and homomorphism, "So that's why I put it together. I just remember using that word ring homomorphism over and over again, so that's why I thought they were connected." Contrary to these results, one student portrayed concept connections in this way: "The main concept of connections is not only based on definitions, but the ways we applied our knowledge of each concept, so for instance, in rings, once we covered the definition of what makes a set a ring, we talked about applications of rings." Thus, these graduate students constructed varying perceptions of concept connections for various reasons.

Implications for Future Research

This exploratory research study resonates with past assumptions that students are not developing accurate concept images and concept definitions of abstract algebra concepts. Professors, especially, should find this research useful since many mathematics professors may not know what the students are actually learning or not learning in their classes. Abstract algebra has historically been a course where there exists a mismatch between what the professor assumes students are learning and what knowledge students are actually attaining. Hence, providing professors a snapshot of what students identified as key concepts (or about which they indicated confusion) would be immensely beneficial.

In future work, I hope to utilize the methodology and results of this study to investigate connections between school algebra and the abstract algebra course. These connections are often missing from classroom instruction, which leads to underdeveloped abstract mathematical thinking. In fact, the participants of this study seemed to be searching for these connections. Heather affirmed:

> Making connections with other courses or ideas, I feel like that it is really hard to do it but it is important and it's helpful. I really wished I knew this before I taught so that I can make better connections in my own teaching. … Because then I would have been able to provide more let's say examples or even provide more opportunities for them to think about things to make connections between the mathematical ideas.

This student wanted to establish concept connections, especially between abstract algebra and school algebra. Likewise, all three students discussed applications (or lack thereof) from abstract algebra. Andrew claimed abstract algebra was not useful due to his inability to apply the theorems and definitions to real-life. Therefore, in order for students to truly construct new algebraic knowledge, there must be connections to past knowledge to successfully internalize and interiorize new ideas. Future research will examine these connections between course concepts and school algebra in order to provide educators a reference to enhance student learning of abstract algebra.

References

Asiala, M., Brown, A., Kleiman, J., & Mathews, D. (1998). The development of students' understanding of permutations and symmetries. *International Journal of Computers for Mathematical Learning*, *3*, 13-43.

Asiala, M., Dubinsky, E., Mathews, D. M., Morics, S., & Oktaç, A. (1997). Development of students' understanding of cosets, normality, and quotient groups. *Journal of Mathematical Behavior, 16*(3), 241–309.

Brown, A., DeVries, D. J., Dubinsky, E., & Thomas, K. (1997). Learning binary operations, groups and subgroups. *Journal of Mathematical Behavior, 16*(3), 187–239.

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 509-535). Thousand Oaks, CA: Sage.

Cook, J. P. (2012). *A guided reinvention of ring, integral domain, and field*. (Doctoral dissertation). Retrieved from ProQuest. (UMI 3517320)

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics, 27*, 267–305.

Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147–164). New York, NY: Macmillan.

Freedman, H. (1983). A way of teaching abstract algebra. *American Mathematical Monthly, 90*(9), 641–644.

Gallian, J. A. (1990). *Contemporary abstract algebra* (2nd ed.). Lexington, MA: D. C. Heath.

Hart, E. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. In J. Kaput & E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning: Preliminary analyses and results* (pp. 49–63). Washington, DC: Mathematical Association of America.

Hazzan, O. (1999). Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics, 40,* 71-90.

Larsen, S. (2004). *Supporting the guided reinvention of the concepts of group and isomorphism: A developmental research project*. Unpublished Doctoral Dissertation, Arizona State University.

Larsen, S., Johnson, E., Rutherford, F., Bartlo, J. (2009). *A local instructional theory for the guided reinvention of the quotient group concept*. Conference on research in undergraduate mathematics education. Raleigh, NC. Retrieved January 4, 2013 from http://sigmaa.maa.org/rume/crume2009/Larsen_LONG.pdf.

Leron, U. & Dubinsky, E. (1994). *Learning abstract algebra with ISETL*. New York, NY: Springer-Verlag.

Leron, U. & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly, 102*(3), 227-242.

Leron, U., Hazzan, O., & Zazkis, R. (1995). Learning group isomorphism: A crossroads of many concepts. *Educational Studies in Mathematics, 29*, 153–174.

Novak, J. D. & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*, (Technical Report IHMC CmapTools 2006-01 Rev 01-2008). Retrieved from Florida Institute for Human and Machine Cognition: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pedersen, J. J. (1972). Sneaking up on a group. *Two Year College Mathematics Journal, 3*(1), 9–12.

Selden, A., & Selden, J. (1987). Errors and misconceptions in college level theorem proving. In *Proceedings of the 2nd international seminar on misconceptions and educational strategies in science and mathematics*. Ithaca, NY: Cornell University.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12,* 151-169.

Taylor, S. J., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings.* New York: Wiley.

Trochim, W. M. K. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning, 12,* 1-16.

Von Glasersfeld, E. (1989). Constructivism in education. In T. Husen & T. N. Postlethwaite, (Eds.), *International encyclopedia of education, Supplement* (Vol.1, pp. 162–163). Oxford, UK: Pergamon Press.

Von Glasersfeld, E. (1991). Abstraction, re-presentation, and reflection: An interpretation of experience and of Piaget's approach. In L. P. Steffe (Ed.), *Epistemological foundations of mathematical experience* (pp. 45–67). New York, NY: Springer.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*, 101–119.

Zazkis, R., & Hazzan, O. (1999). Interviewing in mathematics education research: Choosing the questions. *Journal of Mathematical Behavior*, *17*(4), 429-439.

125

MATHEMATICAL THINKING IN ENGINEERING AND MATHEMATICS STUDENTS

Jenna Tague
The Ohio State University

Key words: Advanced Mathematical Thinking, Cognition, Post-Secondary Education

The past decades have brought a multitude of calls for improving the mathematical education of Science, Technology, Engineering, and Mathematics (STEM) students as well as increasing the number of STEM graduates (Ferrini-Mundy & Güçler, 2009). However, there is a need to examine what mathematics and mathematical thinking is needed for these STEM disciplines. Recent work has shown that the mathematical skills needed are highly specialized (Author) and further work has shown that there are mismatches between *users* of mathematics and *teachers* of mathematics in regard to expectations for students' knowledge (Ferguson, 2012).

There is currently no consensus on how to address the lack of cohesion between the *teachers* and *users* of mathematics or how specialized mathematical skills identified above might be supported in coursework (Wankat, 2008). Devoting attention only to developing new courses accommodating specialized mathematical needs, however, can leave students with gaps in schema development related to discipline needs. Curriculum and instruction design efforts are less likely to be effective in the absence of an understanding of how different STEM audiences think mathematically. This study examined the mathematical thinking of two purposefully selected students (one from mathematics and one from engineering) enrolled at a large Midwestern university as a starting place in addressing this gap. Each student was interviewed and given two open-ended questions and one typical "word problem." Interviews were each approximately 90 minutes long and students were encouraged to explain their reasoning and thinking processes.

Interviews were analyzed through the socio cultural lens of zone theory (Valsiner, 1997) in order to investigate the resources the students drew upon while thinking mathematically. Additionally, a mathematical modeling cycle (Blum & Leiβ, 2007) allowed for cataloguing the particular phases involved in the participants' mathematization processes. Initial findings indicate that the mathematics student was comfortable creating parameters and estimating values whereas the engineering student primary focus was on accuracy. The engineering student also validated not just the numerical estimates he made, but also the solution methods he was considering. Differences and similarities between the mathematical thinking of the two students will be discussed as well as possible instructional implications.

References

Blum, W. & Leiß, D. (2007). How do students' and teachers deal with modelling problems? In: Haines, et al. (Eds), *Mathematical Modelling: Education, Engineering and Economics*. Chichester, UK: Horwood Publishing, 222-231.

Ferrini-Mundy, J. & Güçler, B. (2009). Discipline-based efforts to enhance undergraduate STEM education. New Directions for Teaching and Learning. 117, 55-67.

Valsiner, J. (1997). *Culture and the development of children's actions: A theory of human development* (2nd ed.). Chichester, UK: John Wiley & Sons.

Wankat, P. C. (2008). Pedagogical training and research in engineering education. *Chemical Engineering Education*, 42(4), p. 203-210.

PENCASTS AS EXEMPLARS IN DIFFERENTIAL EQUATIONS

Jenna Tague
The Ohio State University

Jennifer Czocher
Texas State University

Amanda Roble
The Ohio State University

Greg Baker
The Ohio State University

Key words: Smartpens, Cognition, Post-Secondary Education

A substantial amount of students' time in mathematics courses at the undergraduate level is spent working homework problems. Many textbooks are designed to encourage students to look for similar problems while working homework problems (Lithner, 2004). This approach requires very little engagement with the mathematics content.

The context of this study is a differential equations course for engineering students whose goal is to introduce differential equations content through paradigmatic engineering examples and mathematical modeling. To encourage further engagement with the material and to provide asynchronous instructional scaffolding outside the classroom, we sought technology that was dynamic and provided exemplars of mathematical modeling in differential equations.

We chose to create these exemplars using a Livescribe smartpen as an instructional medium. A smartpen is a ballpoint pen with an internal infrared camera and audio recording device. The user writes on "digital paper", which is covered in small dots that locate the pen on the page. The audio is synchronized with written text to create a flash video called a pencast, which can then be shared or embedded into the course website.

Over the past two years, we have iteratively designed ways to assess how students use pencasts and what effect pencasts have on students' knowledge of differential equations. These methods include quantitative and qualitative surveys. Quantitative survey results have shown that students find the pencasts helpful, and self-assessed that they were better able to solve homework problems on their own after watching related pencasts (Roble, Tague, Czocher, & Baker, 2013). Qualitative results indicate that students appreciated the explanation that was provided. Usage survey reports indicated that students used the pencasts to study for test and quizzes which is similar to other reports on pencast usage. However, our students also reported that they used the pencast when solving other similar problems and when they were stuck on homework problems. When comparing pencasts to static solutions, one student remarked it was, "easier to see how problems are solved [via pencast] than the book because the explanation isn't in between each of the steps, it's explained through talking."

In this poster, we will share our full survey results from across semesters, provide demonstrations of how to use a smartpen and show the final flash video product. Additionally, we will share how we chose the smartpen as an instructional medium, and adapted it to our purposes in a differential equations course for engineering students.

References

Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *Journal of Mathematical Behavior*, 23(4), p. 405–427.

Roble, A., Tague, J., Czocher, J., & Baker, G. Pencasts as Exemplars of Mathematical Modelling for Engineering Students. Proceedings of the International Conference on Engineering Education, Marrakesh, Morocco, 2013.

127

THE CONSTRUCTION OF A VIDEO CODING PROTOCOL TO ANALYZE INTERACTIVE INSTRUCTION IN CALCULUS AND CONNECTIONS WITH CONCEPTUAL GAINS

Matthew Thomas
University of Central Arkansas

Instruments called concept inventories are being used to investigate students' conceptual knowledge of topics in STEM fields, including calculus. One interactive instructional style called Interactive-Engagement has been shown to improve students' gains on such instruments in physics. In this paper, we discuss the development of a video coding protocol which was used to analyze the level of Interactive-Engagement in calculus classes and investigate the correlation with gains on the Calculus Concept Inventory.

Key words: Calculus, instruction, interactive teaching, conceptual learning

Conceptual understanding has been a recent area of interest in undergraduate mathematics and in other STEM fields. This interest has manifested in the construction of instruments called concept inventories to measure students' conceptual understanding. Previous studies have found correlations between interactive instructional techniques, particularly one called Interactive-Engagement (IE), and gains on concept inventories. In this study, we investigate possible connections between conceptual gains on one such instrument, the Calculus Concept Inventory, and Interactive-Engaged instruction. To measure IE, we constructed a coding protocol that quantitatively measures the level of IE in a classroom. This study serves two purposes: (1) to develop a coding protocol to quantify IE instruction, and (2) to connect the results of the coding protocol with scores on the CCI.

Background

Conceptual Knowledge

Historically, there has been a division between the teaching of computational and conceptual material (Rittle-Johnson, Siegler, & Alibali, 2001). These "sharply contrasting orientations" (A. G. Thompson, Philipp, T. Thompson, & Boyd, 1994, p. 1) can be seen in the recent "math wars," where proponents of traditional mathematics typically emphasize procedural fluency and proponents of reform-based (or standards-based) mathematics emphasize conceptual understanding (Schoenfeld, 2004).

Rittle-Johnson et al. (2001) define *procedural knowledge* as "the ability to execute action sequences to solve problems" (p. 346). In contrast, *conceptual knowledge* is defined as "implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain" (p. 346). For example, conceptual knowledge might be indicated by a student's understanding of the relationships between algebraic and graphical representations of functions. One way that conceptual knowledge can be demonstrated is by applying known principles or techniques in new situations. For example, recognition of the same topic, such as optimization, in a different subject area or context gives credence to the claim that conceptual understanding has been obtained (Hughes Hallett, 2006, p. 4).

Concept Inventories

Conceptual understanding may be measured through instruments called concept inventories.

Concept inventories are tests designed to measure the most basic knowledge in a field (Epstein, 2007). Typically, the tests are given in a multiple choice format, and involve no computation. When given as a pretest and posttest, the instruments measure the change in conceptual knowledge students undergo during a course. Many studies of conceptual understanding in physics education use concept inventories (e.g. Hake, 1998; Halloun, 1985; Malone, 2008; Rhoads & Roedel, 1999), and other disciplines are using them with increasing frequency (Libarkin, 2008).

The first concept inventory was the Force Concept Inventory (FCI), written by Hestenes, Wells, and Swackhamer (1992) to measure students' conceptual knowledge in introductory mechanics courses. Drawing upon the FCI, Epstein wrote a 22-question concept inventory for introductory calculus in 2007 (Epstein, 2007, 2013).

Interactively-Engaged Instruction

Interactively-Engaged (IE) instruction has been linked to gains in conceptual understanding as measured by concept inventories. IE was defined by Hake (1998) as a collection of methods designed, at least in part, to promote conceptual understanding through "heads-on (always) and hands-on (usually)" (p. 1) activities which lend themselves to immediate feedback through discussion with peers and/or instructors. In this study we operationalize the concept of an IE classroom in a way that allows IE to be quantitatively measured and explore potential correlations between Interactive-Engagement and gains on the CCI.

Previous studies that consider correlations between concept inventory scores and interactive instruction have relied on instructor and/or student self-reporting to quantify levels of IE in classrooms. We eliminated the need for self-reporting by developing a protocol and coding videos ourselves. Our protocol also allows for the examination of IE as a continuum rather than a dichotomous variable, as has been done before. For example, a study by Prather, et al. (2009) relied on instructor self-reporting of interactivity levels, where questions were designed to determine how frequently "interactive learning strategies" (p. 322) were implemented, and how often students made predictions or were asked questions during class. Rhea's (n.d.) study relied on student and instructor reporting of interactivity levels.

Methods

All students taking introductory calculus in the fall semester of 2010 at a large southwestern university took the CCI as a pretest and posttest. Instructors teaching introductory calculus again in the spring semester of 2011 were invited to participate in the study. Of the ten instructors who taught introductory calculus in both Fall 2010 and Spring 2011, five agreed to be videotaped in the classroom three times during the semester. The student scores and instructor videos were collected during different semesters for logistical reasons, however instructors indicated that they were using the same instructional style both semesters. This difference in timing of data collection should be considered when interpreting the results of the study.

Coding Process

We developed a set of interaction types including descriptions of what would constitute each type of interaction. We then used three videos to refine the descriptions of the interactions, develop key examples, and add categories of interactions that were not previously anticipated.

The final coding protocol was applied to the 12 videos not used for the development of the coding protocol. It is important that the results of a video coding protocol are not dependent upon the individuals coding, so that the coding protocol can be used by other researchers to reach similar results. Two researchers independently coded one video from each instructor and created a master code to resolve any disagreements. The independent codes of the two researchers were over 80% reliable for each interaction type in each of 3 videos. The remaining 9 videos were coded by one of the researchers.

CCI Analysis

Traditionally, concept inventory scores are analyzed using a measure called the normalized gain, which is the fraction of gain achieved out of the total that could be obtained, defined as:

$$\langle g \rangle = \frac{\textit{Posttest Score} - \textit{Pretest Score}}{\textit{Maximum Possible Score} - \textit{Pretest Score}} \quad . \tag{1}$$

This score is typically defined at the classroom level, though the effect of computing these scores at the student level has been addressed as well (Bao, 2006; Coletta & Phillips, 2005). Bao found that the differences could largely be attributed to differences between classes where all students gained uniformly and those where the rank order of students changed. This change in rank order might occur in situations where an instructional style is particularly effective for a subset of the population, like students with initially lower ability. We considered the effect of using both student-level and instructor-level normalized gain scores.

Results

Final Coding Protocol

Videos were coded by classifying each interaction. Only interactions around non-routine problems were considered admissible. For the purposes of this study, we considered routine problems to be those that were completely procedural; they required no interpretation and were algorithmic in nature, such as finding the derivatives of a list of functions. In the classrooms observed, wholly procedural problems were uncommon. Most problems included a real-world context or were building towards a discussion of underlying concepts. For example, all related rates problems observed were considered non-routine because they included an interpretation, such as determining how to model the problem or how to interpret a solution in real-world terms. A problem involving a conical sand pile might include a conversation about the shape of a sand pile, or the interpretation of the sign of the rate of change of the radius with respect to time.

The scope of an interaction was determined by the framing of the question or comment which initiates the interaction. For example, an instructor might ask "what is the value of x in this problem?". In this case, the question marks the beginning of the interaction, and the end of the interaction occurs when the value of x is determined. If the instructor instead asked "how would we set this problem up?", the interaction would be considered to conclude when the setup for the problem has been addressed. Though not frequent, this allows for a single interaction to include multiple exchanges and/or multiple students.

All interactions were categorized as either public or private. Private and public interactions may contribute to student gains in different ways, and the literature does not currently distinguish between these types of interactions in an IE classroom. By dividing interactions in this way, we can investigate whether public or private interactions encourage greater gains. Then, both private and public interactions were categorized by the initiator of the interaction. The initiator of an interaction is the person who introduces the content of a conversation. In public conversations, this was very clear, as the instructor typically initiates interactions unless a student specifically asks a question or proposes an idea. The only category of interactions that was not initiator dependent was called Developing Concepts. These episodes consisted of a sustained discussion on the conceptual content on a topic. For example, an instructor might develop the idea of L'Hopital's rule by appealing to notions of derivatives and rate of change to motivate the statement of the rule, or a student might ask whether L'Hopital's rule has anything to do with rates of change.

Public Interactions

In order to be coded as a public interaction, an interaction must be visible and audible to the majority of the class, the content of the interaction must be calculus-based, and must access students' knowledge, not students' perception of their knowledge. Accessing students' perceptions occurred frequently when an instructor asked "does that make sense?". An answer to this question does not provide the instructor with any information about students' understanding, only whether they think they understand. Similarly, choral response questions, where the answer was clear from the instructor's question, almost never provided substantial information to an instructor and were inadmissible. These questions typically only assessed student perception of understanding, and never provided opportunities for discussion to continue. If a choral response question did lead to a substantial conversation, this conversation was eligible to be counted as an interaction.

Public interactions typically took place when a student asked a question or made a suggestion during class by raising their hand. If the student made a suggestion that extended the conversation beyond the scope of the current conversation, this was considered a new, student-initiated interaction, as opposed to a continuation of the occurring interaction. Student-initiated interactions can include incomplete attempts, such as an incompletely formed question or suggestion. For student-initiated interactions, a student attempting to contribute to the discussion was the key factor in identifying the student as the initiator.

Public interactions were further divided by the initiator, and then by type of interaction. The student-initiated interactions consisted of developing strategies, sensemaking, and checking for correctness (see Table 1 for descriptions and examples). These types of interactions were derived from the descriptions of IE classrooms given by Hake (1998) and Epstein (2007).

Table 1: Public Student-Initiated Interaction Categories

Category Name	Description	Examples
Developing strategies	A student suggests or asks a question about how to solve a problem. This may be a suggestion or question specific to the problem at hand or about a class of problems.	Suggesting a new step in a problem, or asking whether a different solution path would be successful.
Sensemaking	A student makes a comment or raises a question about interpreting content in the course.	Interpreting answers, units, magnitudes, or signs of answers in the work being discussed.
Checking for correctness	A student makes a comment which corrects or asks about the correctness of a solution or step in a solution process.	A student asks why a particular step in a process was justified, or points out a mistake.

Instructor-initiated interactions were those in which the instructor specifically asked a question or began an interaction where the instructor determined the topic of the conversation. These interactions were divided into the categories: promotes sensemaking, promotes checks / connections to previous material / extensions beyond current material, encourages revisions from students, check procedures for sense-making, and presentation of problems worked on by students (see Table 2 for descriptions and examples).

Table 2: Public Instructor-Initiated Interaction Categories

Category Name	Description	Examples
Promotes sensemaking	Making a suggestion about how to think about a problem or type of problem.	Drawing attention to notation, such as noting where a parameter is being used in a new way.
Promotes checks / connections to previous material / extensions beyond current material	The instructor extends the discussion outside of the immediate context.	Connecting the immediate material to material that has already been covered or will be covered in the future, referencing a prior problem or prior comment made by a student.
Encourages revisions from students	Explicitly suggesting a revision from the students in the class.	A revision of work the instructor has written himself/herself, or a suggestion to improve upon or work presented by a student.
Check procedures for sense-making	Checking whether the steps of a specific solution process make sense.	Asking why a particular step was done as opposed to a different step, or asking what justifies a particular step of a solution.
Presentation of problems worked on by students	Instructor provides direct and immediate feedback to students immediately after work is completed.	Instructor presents the solution to a problem on the board after students had worked on the problem either individually or in groups, and had completed work on the problem.

Private Interactions and Work Times

Private interactions occurred whenever students were working with each other or discussed content with an instructor when the majority of the class could not hear or see the exchange. When private work time occurred, the number of interactions was counted and the total amount of time students spent actively working was recorded. This time was counted separately depending on whether the private work was groupwork or individual work because groupwork allows students to provide each other with immediate feedback and individual work time provides students opportunities to engage with content. The amount of time devoted to groupwork varied greatly among the five instructors, and has the potential to be another characteristic of an IE classroom. The amount of time in private work was only considered if the private work lasted at least two minutes. Shorter interactions did not allow students to engage with each other sufficiently, or the questions beginning the private work were not of sufficient difficulty to encourage in-depth, conceptual conversations.

In addition to time being provided for groupwork, many of the private work times also included instructor-student interactions as the instructor circulated the room. The number of these interactions was recorded, then further categorized by who initiated the interaction. Instructor-initiated interactions were those in which the instructor asked a specific question of a student, instead of a question that invited conversation but did not initiate discussion of the content.

Miscellaneous (Uncategorized) Interaction Count

A final category was created to capture the interactions which did not fall into any of the other predefined categories. These included interactions where the topic was precalculus material or may not have qualified as any other particular type of interaction.

Results of Coding

The counts of each type of interaction are given in Table 3 along with the associated normalized gain scores. Among these five instructors, the coding protocol distinguished instructional activities in meaningful ways, both by indicating which instructors were more interactive and by quantifying differences between interactive instructors.

Table 3: Counts of Types of Interactions by Instructor

	Instructor	A	B	C	D	E
	Developing concepts	3	1	0	0	0
Student work-time, including private interactions	Groupwork time (seconds)	1872	1249	0	865	0
	Individual work time (seconds)	0	504	0	2939	0
	Total work time (seconds)	1872	1753	0	3804	0
	Instructor initiated private interaction	11	18	0	0	0
	Student-initiated private interactions	16	14	0	38	0
Instructor-initiated public interactions	Promotes checks	6	7	6	5	2
	Encourages revisions from students	4	8	0	5	6
	Promotes sense making	9	5	1	3	3
	Feedback on questions answered by students	22	9	14	2	18
	Problem presented which students have worked on	1	0	4	8	1
Student-initiated public interactions	Student initiated developing strategies	4	2	1	4	4
	Student initiated sensemaking	2	0	3	8	10
	Student initiated check correct	4	2	4	4	10
	Misc. (Uncategorized) interaction count	52	85	40	51	58
	Normalized gain	0.239	0.271	0.190	0.246	0.259

CCI Gains

There were 26 sections of the course, with a maximum capacity of 35 in each section. Most classes were near capacity, and on average 18.5 students per section participated in the study, with a range of 10 to 26 participating students. The classrooms of the 5 instructors who agreed to participate represented a spectrum of normalized gain scores on the CCI ranging from 0.19 to 0.27, near the national average. The mean normalized gain for the entire participant group at the large, southwestern university where our study was conducted was 0.25, meaning that 25% of the previously unknown concepts was learned during the course. Normalized gain scores for the entire 26 sections ranged from 0.14 to 0.36.

Correlations Between Counts and Gains

The total number of classroom interactions (including those considered "Miscellaneous") was significantly related to student gains, as demonstrated in Figure 1 and reported as Model 1 in

Table 4. Among the specific categories, “Encourages Revisions from Students” was significantly related to student gains, shown in Figure 2 and reported as Model 2 in Table 4.

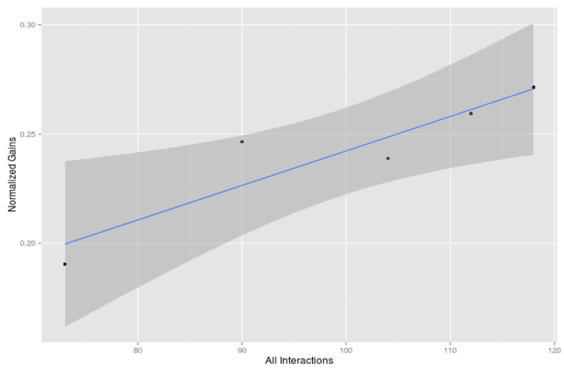


Figure 1: Normalized gains versus all interactions

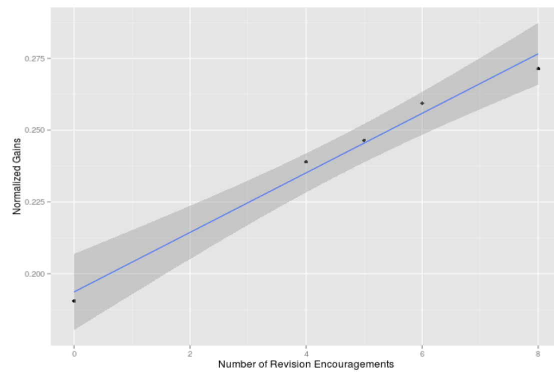


Figure 2: Normalized gains versus number of revisions encouraged

Table 4: Regression Results for Exploratory Analysis

Variable	B	SE(B)	β	*t* (df = 3)	Sig(*p*)	R^2
Model 1: All Interactions						
Constant	0.084	0.039		2.181	0.117	0.849
Interactions	0.002	0.0003	0.922	4.108	0.026	
Model 2: Number of Revisions Encouraged						
Constant	0.194	0.004		46.49	< 0.001	0.983
Interactions	0.010	0.001	0.992	13.22	< 0.001	

Note: B indicates the unstandardized regression coefficient. β indicates the standardized regression coefficient.

Student-Level Analysis of Student Scores

While the previous results were conducted at the classroom level, one can use a statistical technique called Hierarchical Linear Modeling, also known as multi-level modeling, to analyze scores at the student level (Gelman & Hill, 2007). Using this technique, we can analyze student-level gain scores and the results of the video coding protocol at the instructor-level. We calculated normalized gain scores for students using the same formula as was used for classrooms, and constructed a model called a null model which partitions the variance between the student-level and instructor-level. We found that over 99.9% of the variance lies at the student-level, suggesting that, at the university where the study was conducted, nearly all the variation in student-level normalized gain scores can be attributed to differences between students rather than differences between instructors. This suggests that university-level factors, such as department culture, or choice of textbook, may be affecting student gains, and future studies including video analysis of classrooms from multiple universities may provide further insight as to whether this protocol can help understand the relationship between IE instruction and gains in conceptual knowledge. The discrepancy between the instructor-level analysis and the student-level analysis suggests that this relationship is perhaps more complicated than previously thought, and warrants further investigation.

Conclusions

The video coding protocol developed in this study provides a means for analyzing additional classrooms to further investigate the connections between IE instruction and gains in conceptual learning as measured on a concept inventory. When analyzed at the classroom-level, as is traditionally done, our data indicated that despite the small sample size, this coding protocol may describe IE behaviors which are tied to gains on the CCI. While the analysis of the Hierarchical Linear Model suggests that care needs to be taken in interpreting the results at the classroom level, the qualitative differences between classrooms demonstrated by the use of the protocol suggest that this tool can be useful in further investigations of IE instruction.

References

Bao, L. (2006). Theoretical comparisons of average normalized gain calculations. *American Journal of Physics*, *74*(10), 917. doi:10.1119/1.2213632

Coletta, V. P., & Phillips, J. a. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics*, *73*(12), 1172. doi:10.1119/1.2117109

Epstein, J. (2007). Development and validation of the Calculus Concept Inventory. In *Proceedings of the Ninth International Conference on Mathematics Education in a Global Community* (pp. 165–170).

Epstein, J. (2013). The Calculus Concept Inventory - Measurement of the effect of teaching methodology in mathematics. *Notices of the AMS*, *60*(8), 1018–1026.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. doi:10.1119/1.18809

Halloun, I. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, *53*(11), 1043. doi:10.1119/1.14030

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*(3), 141–158. doi:10.1119/1.2343497

Hughes Hallett, D. (2006). What have we learned from calculus reform? The road to conceptual understanding. *MAA NOTES*, *69*(July), 43.

Libarkin, J. (2008). Concept inventories in higher education science. In *National Research Council Promising Practices in Undergraduate STEM Education Workshop 2*. Washington, D. C.

Malone, K. (2008). Correlations among knowledge structures, force concept inventory, and problem-solving behaviors. *Physical Review Special Topics - Physics Education Research*, *4*(2), 1–15. doi:10.1103/PhysRevSTPER.4.020107

Prather, E., Rudolph, A. L., Brissenden, G., & Schlingman, W. M. (2009). A national study assessing the teaching and learning of introductory astronomy. Part I. The effect of

interactive instruction. *American Journal of Physics*, *77*(4), 320. doi:10.1119/1.3065023
Rhea, K. (n.d.). The Calculus Concept Inventory at a large research university. *Unpublished manuscript*.
Rhoads, T. R., & Roedel, R. J. (1999). The Wave Concept Inventory - A cognitive instrument based on Bloom's Taxonomy. In *Proceedings of the 29th ASEE/IEEE Frontiers in Education Conference* (Vol. 3, pp. 14–18). IEEE.
Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362.
Schoenfeld, A. H. (2004). The math wars. *Educational Policy*, *18*(1), 253–286. doi:10.1177/0895904803260042
Thompson, A. G., Philipp, R. A., Thompson, T. W., & Boyd, B. A. (1994). Calculational and conceptual orientations in teaching mathematics. In A. Coxford (Ed.), *1994 Yearbook of the NCTM* (pp. 79–92). Reston, VA: NCTM.

TEACHING METHODS AND STUDENT PERFORMANCE IN CALCULUS I

Lisa Mantini
Oklahoma State University

Barbara Trigalet
Texas Academy of Biomedical Sciences

R. Evan Davis
Oklahoma State University

Classroom teaching in multiple sections of Calculus I at a large comprehensive research university was observed and coded using the Teaching Dimensions Observation Protocol (TDOP). Multiple teaching styles were identified ranging from low engagement to moderate engagement to high engagement sometimes including student group work. Student performance on two course-wide uniform exams and on the Calculus Concept Inventory was analyzed for any correlations with teaching methods. Significant correlations were found with high engagement teaching styles on both the first exam and the final exam. However, no significant correlations were found on the Calculus Concepts Inventory, indicating that students may not have exerted much effort on this assessment.

Key words: [Calculus instruction, classroom observations, student performance, calculus concepts inventory, teaching dimensions observation protocol]

Introduction and Literature Review

The United States is not producing enough graduates in Science, Technology, Engineering and Mathematics (STEM) (Bressoud, 2011) and the need is particularly great in the mathematically intensive majors. However, college freshmen entering one of the STEM majors face a significant hurdle in Calculus I. Currently, the Mathematical Association of America is investigating the teaching of college calculus courses nationwide to “measure the impact of the various characteristics of calculus classes that are believed to influence student success” (Bressoud et al., 2013, p. 2). In order to measure the impact of those characteristics, they must first be defined and described. As Speer, Smith and Horvath note, “research on collegiate teachers’ actual classroom teaching practice is virtually non-existent” (2010, p. 99). According to Bressoud (2012), “the mathematical community does not have research evidence for instructional strategies that work.” This study seeks to contribute to a growing body of research on actual classroom practice, as well as determine possible correlations between actual classroom practices and student achievement in Calculus I.

While much research has found alternatives to lecture such as “inquiry-oriented” or “constructive process” pedagogies to be successful (Ganter, 1999; Rasmussen, Kwon, Allen, Marrongelle & Burtch, 2006), others have found lecture to be effective (Hora & Ferrare, 2013; Saroyan & Snell, 1997) or preferable to students (Ferrini-Mundy & Güçler, 2009; Murray, 1983). This suggests that there is a need for a detailed description of in-class instruction to capture the relations among instructors, students and classroom environments. Porter (2002) notes that careful analysis of teaching can help identify methods that contribute to student achievement.

A growing trend in the assessment of student understanding is the use of Concept Inventories, dating back to the work in physics of Halloun and Hestenes (1985) in developing the Force Concept Inventory (FCI). The FCI is intended to serve as a reproducible and objective measure of how a course improves comprehension of principles (Epstein, 2007); higher gains are seen after interactive engagement pedagogies in which students receive immediate feedback in class on their understanding of a topic. Similarly, the Calculus Concept Inventory (CCI) (Epstein, 2012) measures conceptual understanding of the principles of calculus through the use of multiple choice questions requiring little to no calculation. Typically the CCI is given as a pre-test and post-test in one semester, and sections of the course are compared by comparing their normalized gain, which is the ratio of actual gain in the class average score (post-test mean less pre-test mean) divided by maximum possible gain (maximum possible score less pre-test mean).

Research Questions

The research questions addressed by this study are:

1. What instructional practices including teaching methods, pedagogical moves, instructor/student interactions, cognitive engagement and instructional technology are being used in Calculus I at a large research university?
2. (a) Which of these practices correlate to increased student conceptual understanding as measured by normalized gain on the Calculus Concepts Inventory? (b) Which of these practices correlate to higher average student performance on a uniform final exam?

Research Methodology

In this section we describe the setting of the study, the participants and method of selection, the sampling techniques, the instruments used, and the data collection and analysis.

Setting and Participants. At the large, comprehensive research university during the semester when this study took place, Calculus I was taught in small sections with from 36 to 43 students per section. These sections met for either four 50-minute meetings or three 75-minute meetings per week; class start times ranged from 8:00 AM until 2:30 PM. Section enrollments were unrestricted, and students self-enrolled into their preferred section. Study participants consisted of 10 volunteers from among the section instructors, each teaching 1 or 2 sections of Calculus I. Two instructors were tenured professors with substantial teaching experience; the remaining instructors had held the Ph.D. for four years or less or were advanced doctoral students within a year or two of earning the Ph.D. Two instructors were teaching their own section of Calculus I for the first time; all others had prior experience as an independent instructor in Calculus I. Four instructors were in their first year of teaching at the study institution. Study participants accounted for over 90% of the sections of Calculus I taught during the semester in question, and the 454 students enrolled in these sections accounted for over 90% of the students enrolled in Calculus I during that semester.

The students enrolled in sections taught by participating instructors were asked to volunteer for the study. From among the volunteers, study participants were students who completed the various assessment instruments used. There were 350 student participants who completed both the uniform Exam 1 and common Final Exam, accounting for over 70% of students enrolled in Calculus I that semester. Due to spotty attendance in class on the days when the CCI test was administered, scores on the CCI pre- and post-test are available for 208 students, representing from 45% to 94% of students from each section participating, except for one section with only 34% of students volunteering.

Data Collection. All instructors teaching Calculus I collaborated on writing the uniform Exam 1 and Final Exam. Questions were fairly standard and emphasized calculations but some conceptual questions and some real-world applications were included. Grading was done uniformly, with one instructor grading one problem on all papers. Scores for student study participants were reported to the researchers. The Calculus Concepts Inventory was administered in class early in the semester and late in the semester by all instructors. This is a multiple choice instrument requiring little calculation which tests conceptual understanding of calculus concepts (Epstein, 2012). Instructors scored their own sections and reported results to the researchers.

Classroom observations were conducted using the Teaching Dimensions Observation Protocol (TDOP) created by M. Hora and J. Ferrare (Hora and Ferrare, 2010). This instrument codes which of multiple behaviors by teachers or students are observed during each 2-minute interval of an observation. It has been used previously to classify instructional behaviors in college-level instruction in Calculus (Code, Kohler, Piccolo, and MacLean, 2012) and across disciplines (Hora and Ferrare, 2013). Instructors participating in the study had access to the instrument and were aware that the broad categories being observed were Teaching Methods, Pedagogical Moves, Instructor-Student Interaction, Cognitive Engagement, and Instructional Technology (see Appendix for table listing specific codes). Each section in the study was observed 3 times except for one section which was only observed twice. Before each observation, the observer contacted each instructor to ascertain that the observed class period would be what the instructor would call "typical."

Other observation instruments were considered and rejected for this study. Among these were the Teacher Behavior Inventory (TBI) (Murray, 1983), which gathers subjective accounts from students assessing instructor behaviors, and the Reformed Teaching Observation Protocol (RTOP) (Sawada et al, 2002), which aims to evaluate the extent to which instruction meets the goals of being inquiry-oriented or student-centered, and thus does not provide a descriptive account of teaching behaviors (Hora and Ferrare, 2013).

Before using the observation instrument, the researchers observed videotaped instruction and coded together, in order to train themselves in using the instrument live and to increase observational reliability. All coding was done in person during this study so no video recordings were used, and all observations were performed by one researcher only.

Data Analysis. For each of the 11 sections in the study, observational data from the TDOP were converted into a sequence of 0's and 1's, where a 1 was recorded if that particular behavior was observed in a two-minute interval and a 0 if not. These data were entered into an Excel spreadsheet. Each section was observed 2-3 times, so the total number of observed 2-minute intervals ranged from 50 to 114 per section (some sections met for 75 minutes). We then determined the proportion of observed 2-minute intervals in which each particular TDOP code was observed. This gave us a range of proportions for each TDOP code indicating its relative frequency of use among study participants. Many codes varied little across sections, but those codes that had high variability across sections were noted.

For the initial phase of analysis, student performance was averaged in each section, producing four data points summarizing student performance: the exam 1 average, the final exam average, the sum of the exam averages, and the CCI net gain. CCI net gain is computed as the ratio of the actual section mean gain (post-test mean less the pre-test mean) to the maximum possible mean gain (maximum score minus the pre-test mean). Pearson correlation coefficients were computed between each of the four student performance indicators and TDOP proportions across sections. TDOP categories showing a significant correlation with student performance were analyzed further, individually or in combination. The significant TDOP codes were used to characterize the observed instructional profiles into three categories: low engagement, moderate engagement, and high engagement.

For the final phase, a spreadsheet was created containing the four scores available for each individual student and a number from 1 to 3 indicating the classification of their section instructor's observed instructional profile as low, moderate or high engagement. Additional analysis was performed including ANOVA and ANCOVA to determine if any correlation was present between the instructional profile and student performance.

Research Results

Teaching Practices. Initial findings from the TDOP regarding Teaching Methods indicate that all instructors employ lecturing with visuals, seen in 90% of the two-minute intervals coded. The instructional technique of having students work at their desks, either in small groups (SGW) or by themselves (DW), was observed 11% of the time but it was used by only four instructors, ranging from 16% to 32% of the time in those sections. Several codes in both the Instructor-Student Interaction category and the Cognitive Engagement category varied significantly. Overall, approximately 60% of time intervals coded contained questions asked of the students by the instructors, with students responding in more than 50% of the time intervals coded. However, some instructors used display questions (DQ), asking students to display content knowledge, as often as 85% of the time, others as little as 11% of the time. Among Instructional Technology, the most predominant tool was the chalk board or white board, used 77% of the time. The use of power point slides and a digital tablet varied significantly, ranging from no use to use more than 30% of the time. Other instructional technologies were observed well less than 10% of the time.

The codes SGW and DW, along with ART and some types of questioning, showed a positive correlation with at least one of the student performance measures. Based on this preliminary analysis, we combined some TDOP categories in order to create aggregate codes to compare to the student performance measures. These aggregate codes are described in the following Figure. We determined the proportion of instructional time each of these aggregate codes appeared, and we used these proportions to create an instructional profile for each section in the study.

TMTH	Coded as 1 when any teaching method is coded
SWK	Coded as 1 when either SGW or DW is coded (*students working*)
SVB	Coded as 1 when TMTH = 1, SWK = 0, and any of SNQ, SCQ, SR, ART, RMF, PS, CR or CN is coded (*students verbalizing*)
SENG	Coded as 1 when TMTH = 1 and either SWK = 1 or SVB = 1 (*students engaged)*
LNWV	Coded as 1 when TMTH = 1 and SENG = 0 (*lecture, no student work or verbalization*)

Figure 1: Aggregate TDOP Codes

The resulting instructional profiles indicate a range of instructor behaviors. Code SWK ranged from 0% to 32%; code SVB ranged from 19% to 81% of the time; and SENG ranged from 22% to 88%. As a result, LNWV ranged from 12% to 78%. Note that SWK + SVB = SENG and SENG + LNWV = 100%. The data suggested sorting instructor profiles into three groups. It is notable that one instructor in this section taught two sections, and the two sections were assigned to different instructional profiles.

Profile Name	**SENG range of values**	**Sections observed**
Low Engagement	22% - 40%	4
Moderate Engagement	50% - 70%	3
High Engagement	80% - 88%	4

Figure 2 illustrates the range of instructional profiles seen in the spring observations. The teaching profile bars are arranged from left to right by increasing average score on the final exam and illustrate a statistically significant correlation with the high engagement instructional profile, as discussed in the next paragraph.

Statistical Results for Section Averages. Using our data from the 11 course sections in the study, we calculated Pearson's correlation coefficients between our aggregate TDOP codes SWK, SVB, and SENG and average student scores in each section on four measures: Exam 1, the Final Exam, the Sum of Exam 1 and the Final Exam, and the section's net gain on the CCI. Results indicated that the Exam 1 average was correlated significantly ($p < .05$) with code SENG. The Final Exam average and the Exam Average Sum were both correlated significantly ($p < .05$) with code SWK and correlated highly significantly ($p < .01$) with the combined code SENG.

During this study the course wide net gain on the CCI from the pre-test to the post-test across all sections was only 10.15%, with 5 sections having gains from 3.4-9.6%, 5 sections having gains from 12-15%, and one section having a gain of 20.3%. The section CCI net gain was correlated significantly ($p < .02$) with the code SWK but not correlated with either SENG or SVB.

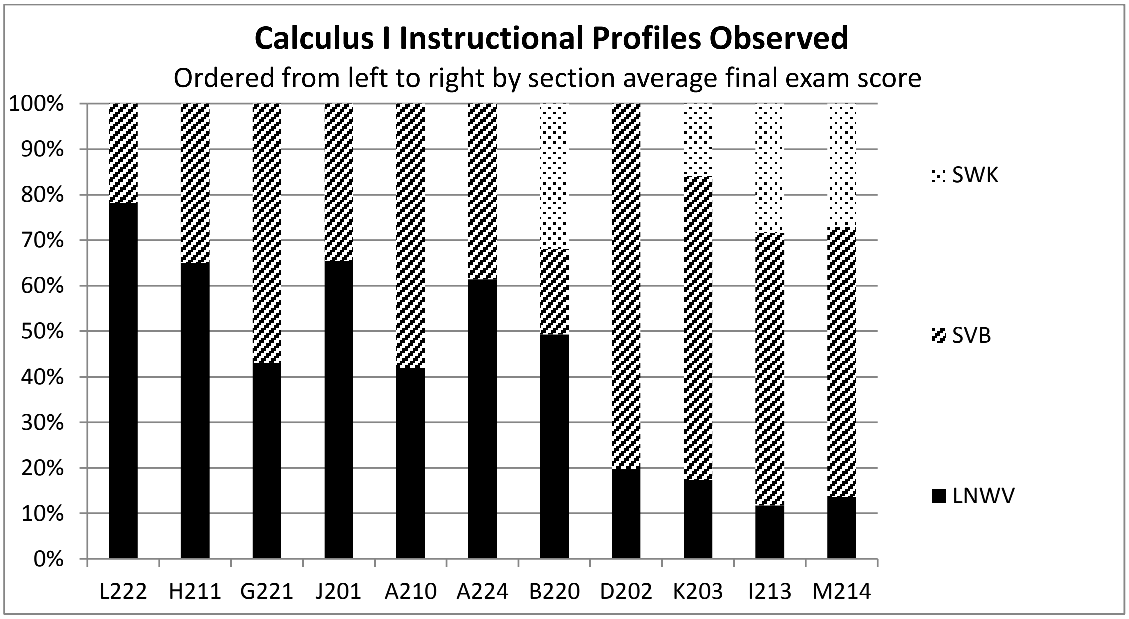


Figure 2: Instructional Profiles

Statistical Results for Individual Student Performance. The student data consisted of scores on each of four assessment measures along with a variable which sorted student scores into three groups according to the instructional profile assigned to their instructor, with 1 indicating the low engagement profile, 2 indicating moderate engagement, and 3 indicating high engagement. The assessment measures considered were the CCI Pre-test, administered in week 1; the score on Exam 1, administered in week 5, the CCI Post-test, administered during week 15; and the uniform Final Exam, administered during week 16 of the semester. We used SPSS software to search for any significant correlations of student performance on the various assessments with the instructional profiles assigned.

We found no significant difference [$F(2,305)=1.88$, n.s.] among students across the three groups in the analysis of the CCI pre-test.

When comparing the scores on common Exam 1, analysis of variance revealed a significant difference in performance [$F(2,347) = 12.84$, $p < .01$] among students. Examination of paired comparisons (Tukey and Scheffe) showed that, while the moderate and low engagement groups did not significantly differ from each other, the high engagement group scored significantly better on the first exam than either of the other two conditions.

When comparing the scores on the common Final Exam, analysis of variance again revealed a significant difference across the three groups [$F(2, 347) = 7.46$, $p < .01$]. Paired comparison between the three groups revealed that, while the difference between the high and low engagement groups was still significant, the difference between the moderate engagement group and either the low engagement group or the high engagement group was not statistically significant. The results for the final exam are interesting in that they suggest that the moderate engagement group "gained ground" on the high engagement group between the first common exam and the final exam, with a higher estimated marginal mean for the moderate engagement group as compared to the high engagement and low engagement groups in an ANCOVA analysis with the final exam as our dependent variable and the first exam as covariate.

We also found no significant difference [$F(2,215)=.08$, n.s.] among students across the three groups in the analysis of the CCI Post-test.

Discussion

Regarding our first research question, we found that the teaching methods observed relied primarily on lecture methods, seen from 68% to 100% of the time. Within lecture methods, though, the use of questioning and other engagement techniques varied significantly. Our data seem to indicate a possible definition of high engagement instruction, but further research is needed. It is interesting to note that all instructional profile groupings included instructors of varying experience levels and both Americans and internationals. However, all high engagement instructional profiles occurred in classes whose start times were between 8:00-10:30 AM.

Regarding our second research question, the correlation of section-wide net gain on the CCI with code SWK agrees with some prior results reported in the literature (Epstein, 2012) but seems suspect. The number of students participating in both CCI pre-test and post-test (n=208) is small and may contain the better students in each section, since many of those absent on the days when the CCI was administered may have been weaker students. The lack of correlation between CCI pre-test and post-test scores and any TDOP variables or other assessments bears further investigation but may indicate a lack of effort by participants on the CCI, which did not count towards their course grade. The high correlation of exam scores with the level of engagement in the instructional profile is very interesting and also deserves further study. This result may imply that there would be a benefit derived from providing training to new Calculus I instructors in questioning techniques and the use of group work.

Further research is desirable to investigate if there is any correlation between teaching methods and persistence in the calculus sequence or student performance in later courses. More observational data might provide richer descriptions of teaching styles in use in Calculus I and further evidence to support the correlations we found. Interviews with instructors might shed light on their decisions with regard to engagement levels and could be relevant to instructional training programs.

References

Bressoud, D. (2011, February 1) Status of the math-intensive majors. [Web log post]. Retrieved from http://www.macalester.edu/~bressoud/pub/launchings/launchings_02_11/launchings_02_11.html#anchor1

Bressoud, D. (2012, July 1) Learning from the physicists. [Web log post]. Retrieved from http://launchings.blogspot.com/2012/07/learning-from-physicists.html

Bressoud, D., Carlson, M., Mesa, V., & Rasmussen, C. (2013). The calculus student: Insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology, 44,* 1-15.

Code, W., Kohler, D., Piccolo, C., & MacLean, M. (2012). Teaching methods comparison in a large introductory calculus class. In S. Brown, S. Larsen, K. Marrongelle, & M.

Oehrtman, (Eds.) *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education*, (pp. 375-379).

Epstein, J. (2007). Development and validation of the calculus concept inventory. In *Proceedings of the Ninth International Conference on Mathematics Education in a Global Community*, (pp. 165-170).

Epstein, J. (2012). *The calculus concept inventory – Measurement of the effect of teaching methodology in mathematics*. Unpublished manuscript, Department of Mathematics, Polytechnic Institute at New York University, Brooklyn, NY.

Ferrini-Mundy, J., & Güçler, B. (2009). Discipline-based efforts to enhance undergraduate STEM education. *New Directions For Teaching And Learning*, *117*, 55-67.

Ganter, S. L. (1999). An evaluation of calculus reform: A preliminary report of a national study. In B. Gold, S. Keith, and W. Marion (Eds.), *Assessment Practices in Undergraduate Mathematics*, (pp. 233-236). Washington, DC: Mathematical Association of America Notes.

Halloun, I. & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, *53*(11), 1043-1048.

Hora, M., & Ferrare, J.. (2010). The Teaching Dimensions Observation Protocol (TDOP). Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Hora, M. T. & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of undergraduate math and science course planning and classroom teaching. *Journal of the Learning Sciences*, *22*(2), 212-257.

Murray, H.G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology*, *75*, 138-149.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, *31*(7), 3-14.

Rasmussen, C., Kwon, O. N., Allen, K., Marrongelle, K. & Burtch, M. (2006). Capitalizing on advances in mathematics and K-12 mathematics education in undergraduate mathematics: An inquiry-oriented approach to differential equations. *Asia Pacific Education Review*, *7*(1), 85-93.

Saroyan, A. & Snell, L.S. (1997). Variations in lecturing styles. *Higher Education*, *33*, 85-104.

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R. and Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics*, *102*, 245-253.

Speer, N. M., Smith III, J. P. & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99-114.

Appendix: TDOP Codes

This table contains all of the codes used in the TDOP.

	Teaching Methods		**Pedagogical Moves**
L	Lecture, no visuals	MOV	Moves into audience
LPV	Lecture, pre-made visuals	HUM	Humor
LHV	Lecture, handwritten visuals	RDS	Reads verbatim from notes or text
LDEM	Lecture with demonstration	IL	Illustration from real world
LINT	Interactive lecture	ORG	Organization
SGW	Small group work	EMP	Emphasis
DW	Desk work	A	Assessment
CD	Class discussion	AT	Administrative task

MM	Multimedia		**Instructional Technology**
SP	Student presentation	PO	Posters used
	Instructor/Student Interaction	B	Books used
RQ	Instructor rhetorical question	N	Lecture notes actively used
DQ	Instructor display question	P	Pointer used
CQ	Instructor comprehension quest.	CB	Chalk board or white board used
SNQ	Student novel question	OP	Overhead or transparencies used
SCQ	Student comprehension quest.	PP	Powerpoint or digital slides used
SR	Student response	CL	Clickers used
	Cognitive Engagement	D	Demonstration equipment used
ART	Articulation by students	DT	Digital tablet or document camera used
RMF	Reciting or memorizing facts	M	Movie, documentary, other video clip
PS	Problem solving	SI	Simulation
CR	Students create their own ideas	WEB	Website or other online resource
CN	Connections to real world	OB	Object used as part of instruction

129

SLOPE AND DERIVATIVE: CALCULUS STUDENTS' UNDERSTANDING OF RATES OF CHANGE

Jen Tyne
University of Maine

Abstract: *Studies have shown that students have difficulty with the concepts of slope and derivative, especially in the case of real-life contexts. I used a written survey to collect data from 75 differential calculus students. Students answered questions about linear and nonlinear relationships and interpretations of slope and derivative. My analysis focused on students' understanding of slope as a constant rate of change and derivative as an instantaneous rate of change, and what these meant in the context of the problems. Preliminary results indicate that students have more success with slope questions than derivative questions (McNemar's test, $p<0.03$), and that while students correctly use the slope of a linear relationship to make predictions, they do not demonstrate an understanding of the derivative as an instantaneous rate of change and an estimate of the marginal change. Plans for a modified survey and interviews are in place for fall 2013.*

Keywords: *Calculus, Derivative, Rate of Change, Slope, Student Understanding*

Introduction and Research Questions

Robust understanding of derivatives and instantaneous rates of change in calculus requires an understanding of slopes and average rates of change from precalculus (Hackworth, 1994). It is important for the mathematics community to be alert to students' understanding of slope coming into calculus, and to design instruction that expands on that knowledge in teaching the derivative. Students may not have the robust understanding of slope and rates of change that instructors assume, which has consequences for their learning of calculus. Furthermore, calculus students must understand not only instantaneous rates of change, but also continuously changing rates. This covariational reasoning is essential for interpreting dynamic situations surrounding functions (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002).

My study investigates the interpretation and use of slope and derivative in real life contexts. Such applications require students to translate from the context to the abstract level of calculus and then back to the context, a process that requires conceptual knowledge (White & Mitchelmore, 1996). Mathematics educators have emphasized the utility of these sorts of problems, noting that "Not only do real-world situations provide meaningful opportunities for students to develop their understanding of mathematics, they also provide opportunities for students to communicate their understanding of mathematics" (Stump, 2001, p. 88). My focus is on linear and non-linear, one-variable relationships, concepts that should be familiar to first-year calculus students.

The study builds on research about student understanding of slope and rate of change (Stump, 2001; Orton 1984; Barr, 1981; Barr 1980); student understanding of rates of change involving and not involving time (Stump, 2001); student understanding of derivatives (Ferrini-Mundy & Graham, 1994; Bezuidenhout, 1998; Zandieh, 2000); student understanding of the rate of change of linear and non-linear functions (Orton, 1983); and how students' knowledge of rates of change affect their conceptual knowledge of the derivative (Hackworth, 1994). Findings from these studies indicate that students have difficulty understanding slope as a constant rate of change and derivative as an instantaneous rate of change. However, there has not been much research on students' verbal interpretation of the derivative as a rate of change, students' verbal interpretation of slope as a constant rate of change, or students' understanding of the differences in making predictions involving constant and instantaneous rates of change. Since the calculus concepts encountered outside the math classroom have real life contexts, it is important that students are able to interpret these situations and rates of change correctly. The specific research questions are:

- Is there a relationship between calculus students' understanding of slope and their understanding of derivative? Specifically, do students' abilities to interpret the slope as a constant rate of change make them more likely to be able to interpret the derivative as an instantaneous rate of change?
- Do students correctly use the slope and derivative to make valid predictions from models?

Methodology

I collected written solutions to questions from 75 students enrolled in a first semester calculus course at a research-focused university. Students had completed approximately 80% of the one-semester course. My research approach, an analysis of student understanding gained from direct students responses, is consistent with a cognitive theoretical perspective and is well established in the math education community (Byrnes, 2000; Siegler, 2003). This cognitive approach grew out of the need to move away from a product-driven method (looking for just the right answer) to the process-driven approach of cognitive scientists (Schoenfeld, 1987).

While the full survey covered multiple concepts surrounding linear and non-linear one-variable functions (interpreting function values, interpreting slope and derivative values, slope and derivative units, and predicting change and function values), this paper focuses on a subset of the questions, namely predicting change and function values for linear and nonlinear relationships given information about the slope and derivative (see Figure 1).

The function P(t) is the population of a country, in millions of people, where t is the number of years after 2000.

a. P(t) = 2t + 30

i. What does the model P(t) predict the change in population will be from the start of 2020 to the start of 2021? Explain how you got your answer.

ii. A classmate of yours says that the population of the country in the year 2030 would be 90 million. Has your classmate made a valid prediction? Explain why or why not.

b. Now, assuming P(t) is a nonlinear function and $P'(20) = 2$.

i. What does the model P(t) predict the change in population will be from the start of 2020 to the start of 2021? Explain how you got your answer.

ii. A classmate of yours says that the population of the country in the year 2030 would be 90 million, given that P(20)=70 and $P'(20) = 2$. Has your classmate made a valid prediction? Explain why or why not.

Figure 1. A subset of the questions on the calculus survey instrument.

I created questions that were not mechanical in nature and therefore did not assess procedural knowledge; rather, I designed tasks about students' interpretation of slope and derivative to uncover their conceptual knowledge about these topics. The questions about linear relationships are posed to gain understanding of students' knowledge predictions based on linear change. These questions were adapted from a general education textbook written to emphasize conceptual over procedural learning (Franzosa & Tyne, 2010). To answer these questions, students must understand linear change as a constant rate of change. The questions about nonlinear relationships are more complex, and were influenced by *Calculus, 6th edition* (Hughes-Hallet et al., 2013). To answer these questions, students must understand the derivative as an instantaneous rate of change that can be used to predict marginal change, and that the derivative *cannot* be used to make predictions at other input values. Data analysis methods are presented in conjunction with the results below.

Preliminary Results

To determine whether students' answers were correct on the two problems that ask to predict the change in population from 2020 to 2021 (parts (i) in Figure 1), I coded answers as 'correct' and 'incorrect'. The correct response is '2 million' or '2 million people'; anything else (e.g. 72 million, 70 million, or leaving it blank) was considered an incorrect response. I didn't look at the explanations for the purposes of coding. I recorded the combinations of right/wrong in the following 2x2 contingency table (Table 1).

Table 1. Predicted Change in Population from 2020 to 2021

		Nonlinear (2 million people)		
		Right	Wrong	Total
Linear (2 million people)	Right	64%	22%	86%
	Wrong	7%	7%	14%
	Total	71%	29%	N=58

To test the null hypothesis that the probability of getting the linear problem correct is the same as the probability of getting the nonlinear problem correct, I performed McNemar's test (α=0.05), concluding that there was a significant difference in the probabilities (*p*=0.0291). McNemar's test analyzes the right/wrong and wrong/right cells (in particular, the students who answer correctly for the linear question then incorrectly for the non-linear question, as well as the students who answered incorrectly on linear and then correctly on non-linear). If the null hypothesis were true, we would expect these percentages to be similar. My results are not entirely surprising, since one would expect students to be able to interpret slope with more success than interpreting the derivative. Only 7% of the students were more successful interpreting the derivative than the slope (that is, they answered the linear problem incorrectly and the nonlinear problem correctly).

A limitation of this task is that written answers give limited insight into what students truly understand. In some cases, students wrote "2 [million people] because the derivative is 2." Further, as we see later, the majority used the instantaneous rate of change incorrectly to predict the population in 2030. Other researchers have found that depending on the crafting of the question, students sometimes give correct answers for wrong reasons, which makes it difficult to detect misconceptions (Bezuidenhout, 1998). Slightly modified questions and interviews, slated for fall 2013, will investigate student understanding of the instantaneous rate of change as an estimate of the marginal change.

Questions ii (Figure 1) were coded as follows: a response was recorded as 'correct' for part a-ii if the student answered 'yes' and 'incorrect' if the student answered 'no'. Part b-ii was coded 'correct' if the student answered 'no' and coded 'incorrect' if the student answered 'yes'. To answer a-ii correctly, students needed to predict the 2030 populations by using the linear function (plugging in t = 30). For the nonlinear problem (problem b-ii), students needed to understand the prediction was not valid because the instantaneous rate of change at 2020 couldn't be used to predict the change in population from 2020 to 2030. Results are summarized in Table 2.

Table 2. Prediction by Classmate for Population Estimate in 2030

Nonlinear (invalid prediction)				
Linear (valid prediction)		Right	Wrong	Total
	Right	28%	64%	92%
	Wrong	0%	9%	9%
	Total	28%	73%	N=69

I performed the McNemar's test on the 2x2 contingency table to test the null hypothesis that the probability of getting the linear problem correct is the same as the probability of getting the nonlinear problem correct, and concluded that there was a significant difference in the results ($p < 0.001$). Therefore, the distribution of correct responses is different for the two questions, which may highlight some of students' misconceptions about instantaneous change. No students answered the linear problem incorrectly and the nonlinear problem correctly. Further, 64% of the students answered the linear problem correctly and then went on to answer the nonlinear problem incorrectly, often stating that the derivative can be used incorrectly to predict change at different input values (in other words, they assumed that the derivative could be used as a constant rate of change).

An additional indication that students have difficulty interpreting derivatives is that there were more students who left the nonlinear problem blank than students who left the linear problem blank. I intend to use interviews to delve into this issue more. More research is necessary as to the actual student thinking surrounding the non-linear problem; considering the number of students who correctly answered the change in population part but went on to get the 2030 prediction incorrect, it seems that many might not understand the difference between using a derivative for a marginal change, and using a derivative to predict change far in the future.

In summary, my preliminary findings support others' claims that students must have a clear understanding of a constant rate of change in order to understand instantaneous rates of change (Hackworth, 1994). In particular, it was unlikely for students to answer a derivative question correctly (Figure 1, part b) after answering a constant rate of change question incorrectly (Figure 1, part a). While we know that students must understand rates of change to succeed in calculus (Hackworth, 1994), this research adds to the body that shows that rates of change are not well-understood by first-year students, many who have fundamental misconceptions (Bezuidenhout, 1998).

Plans for research in fall 2013 include re-administering the survey to calculus students with slightly modified questions, and interviewing students about their understanding of rates of change, marginal change, predictions, and interpretations.

Discussion Questions for Audience

- What other kinds of questions could be effective in uncovering students' understanding of slope as constant rate of change and derivative as instantaneous rate of change?
- How might the interviews be structured to get at student difficulties in using derivatives and slopes to make predictions?
- Based on the study's research question, in what ways could the research design be improved for future iterations?

References

Barr, G. (1980). Graphs, Gradients and Intercepts. *Mathematics in School*, *9*(1), 5-6.

Barr, G., (1981). Some student ideas on the concept of gradient. *Mathematics in School 10(1),* 14-17.

Bezuidenhout, J. (1998). First-year university students' understanding of rate of change. *International Journal of Mathematical Education in Science and Technology*, *29*(3), 389-399.

Byrnes, J.B. (2000). *Cognitive Development and Learning in Instructional Contexts* (2nd Edition). Pearson Allyn & Bacon.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 352-378.

Ferrini-Mundy, J., & Graham, K. (1994). Research in Calculus Learning: Understanding of Limits, Derivatives, and Integrals. In J. J. Kaput & E. Dubinsky (Eds.), *Research Issues in Undergraduate Mathematics Learning: Preliminary Analysis and Results* (Vol. 33, pp. 29-45). Washington, DC: The Mathematical Association of America.

Franzosa, R. & Tyne, J. (2010). *Algebraic Models in Our World*. Dubuque, IA: Kendall Hunt

Hackworth, J. A. (1994). *Calculus Student's Understanding of Rate* (Doctoral dissertation, San Diego State University).

Hughes-Hallett, D., McCallum, W., Gleason, A., Connally, E., Lomen, D., Quinney, D., …Tucker, T., (2013) *Calculus 6th Edition*. Hoboken, NJ: Wiley

Orton, A. (1983). Students' understanding of differentiation. *Educational Studies in Mathematics*, *14*(3), 235-250.

Orton, A. (1984). Understanding rate of change. *Mathematics in school*, *13*(5), 23-26.

Schoenfeld, A. H. (Ed.). (1987). *Cognitive science and mathematics education*. Lawrence Erlbaum.

Siegler, R. S. (2003). Implications of cognitive science research for mathematics education. *A research companion to principles and standards for school mathematics*, 219-233.

Stump, S. L. (2001). High school precalculus students' understanding of slope as measure. *School Science and Mathematics*, *101*(2), 81-89.

White, P., & Mitchelmore, M. (1996). Conceptual knowledge in introductory calculus. *Journal for Research in Mathematics Education*, 79-95.

Zandieh, M. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. *Research in Collegiate Mathematics Education. IV. CBMS Issues in Mathematics Education*, 103-127.

AN ORIGIN OF PRESCRIPTIONS FOR OUR MATHEMATICAL REASONING

Yusuke Uegatani
Research Fellow of the Japan Society for the Promotion of Science (Hiroshima University)

To build a supplementary theory from which we can derive a practical way of fostering inquiring minds in mathematics, this paper proposes a theoretical perspective that is compatible with existing ideas in mathematics education (radical constructivism, social constructivism, APOS theory, David Tall's framework, the framework of embodied cognition, new materialist ontologies). We focus on the fact that descriptive and prescriptive statements can be treated simultaneously, and consider both descriptive and exemplary models in our minds. This indicates that descriptive statements in mathematics come from our descriptions of models, and prescriptive statements come from the exemplarity of exemplary models. As a practical suggestion from the proposed perspective, we point out that careful communication is needed so that inquiring minds do not recognize the refutation of their arguments as a denial of their way of mathematical thinking.

Key words: Prescriptive perspective, Mathematical reasoning, Theoretical discussion

Introduction

Inquiring minds in mathematics seem to come from the belief that mathematical truth is, in some sense, absolute. Goldin (2003) pointed out the importance to mathematics education of commitment to the integrity of mathematical knowledge. This commitment is of particular importance in undergraduate mathematics education. Consider the belief that the discoverability of new mathematical results is open to everyone, because they do not depend on historical contingency, temporary human discourse, or, especially, on authorities. Undergraduate students will not be willing to continue studying mathematics without this *discoverability belief*. It is useful to identify both the origin of the discoverability belief and why it influences students. However, existing research perspectives on mathematical knowledge (radical constructivism, e.g., von Glasersfeld, 1995; social constructivism, e.g., Ernest, 1998) and mathematical cognition (APOS theory, e.g., Dubinsky & McDonald, 2002; three worlds of mathematics, e.g., Tall, 2008) do not explain how the discoverability belief, or its counterpart in each theory, arises.

One approach to identifying the origin of the discoverability belief is to discuss the ontological aspects of mathematical entities. For example, theoretical frameworks based on the broader interpretation of embodiment (e.g., Lakoff & Núñez, 2000; de Freitas and Sinclair, 2013) can describe how mathematical concepts arise from the physical world. However, they do not have a particular interest in why the consistency of the physical world makes mathematical results consistent.

The above existing research provides explanations for broader educational phenomena. Their scant attention to the discoverability belief is thus unimportant. On the other hand, a supplementary theory from which we can derive practical suggestions for fostering inquiring minds in mathematics is useful. Thus, this paper attempts to build a model to explain the origin of the discoverability belief that supplements existing perspectives.

Duality of Prescription and Description

Ernest (1998) pointed out the limitations of *prescriptive* accounts of mathematics:

> Absolutist philosophies of mathematics such as logicism, formalism, and intuitionism attempt to provide *prescriptive* accounts of the nature of mathematics. Such accounts are programmatic, legislating how mathematics should be understood, rather than providing accurately *descriptive* accounts of the nature of mathematics. Thus they are

failing to account for mathematics as it is, in the hope of fulfilling their vision of how it should be. (pp. 50-51, italics in the original)

Thus, Ernest's (1998) social constructivism takes a descriptive stance. It provides no account of which way of doing mathematics is correct, but rather describes how people do mathematics. Other existing research perspectives for mathematics education also take descriptive stances. They provide no account of which method of understanding mathematics is correct, but merely explain how students do mathematics. However, the preceding discussion is based on the following implicit assumption: we must exclusively choose prescriptive or descriptive philosophies. Both the prescriptive statement "X should be Y" and the descriptive statement "X is Y" can be simultaneously correct.

For example, consider a group $(G,*)$. Suppose that G is a set, and that $*$ is a binary operation on G. The group axioms are as follows: (i) For all a,b in G, $a*b$ is also in G. (ii) For all a,b and c in G, $(a*b)*c = a*(b*c)$. (iii) There exists an element e in G such that, for every element a in G, the equation $a*e = e*a = a$ holds. (iv) For each a in G, there exists an element b in G such that $a*b = b*a = e$, where e is the element defined in axiom (iii). From these axioms, we can derive the statement that the element e postulated in (iii) is unique, and we will say that e postulated in (iii) should be unique if someone argues that there are many elements postulated in (iii). In this case, both statements (involving "is" and "should be") appear correct. This is explained by distinguishing between *in* and *out of* the axiomatic system. The statement that the element e postulated in (iii) is unique is a description of components *in* the system. The statement that the element e postulated in (iii) should be unique (or, more strictly, the statement that we should argue that e postulated in (iii) is unique) is a prescription for us who are *out of* the system. It is important that the element e (or the entity in the system) is not itself bound by the rules of logic, but that all thinking subjects who are out of the system and agree on the group axioms have an obligation to obey some logical inference rules.

In general, a descriptive statement in an axiomatic system and the corresponding prescriptive statement out of the system can be simultaneously correct, because we can always distinguish between in and out of the given system. It is, therefore, an unjustifiable assumption that we cannot simultaneously consider both prescription and description. If we have the ability to self-reflect, and to distinguish between the outside of an axiomatic system and the overall framework that contains the inside and the outside of the system, then prescriptive statements and descriptive statements are dual properties of the overall framework (Figure 1). In addition, it is also important that humans out of the system are prescribed, and the entities in the system are described.

Origins of Prescription

If our reasoning always followed the rules of formal logic, the discoverability belief would be justified by the independence between these rules and human minds. In general, it is difficult to describe the actual practices of mathematics by formal logic (e.g., Fallis, 2003). Thus, we argue that the schemata of descriptions actually prescribe human reasoning.

The schema of descriptions is, for example, the format of implication statements "$P \to Q$." We do not assume that it pre-existed the modus ponens. Rather, we argue that modus ponens pre-existed the schema $P \to Q$, and that the schema was invented to describe a situation where one may infer Q after knowing that P is true. Given the propositions P and $P \to Q$, we usually deduce proposition Q for any propositions P and Q. This does not imply the validity of modus ponens, but implies that there can be a situation where one may infer Q after knowing that P is true. Similarly, the rule of conjecture elimination (inferring

P from $P \wedge Q$) pre-existed the schema $P \wedge Q$, and the rule of universal instantiation (inferring $A(a)$ for any element a from $\forall x\, A(x)$) pre-existed the schema $\forall x\, A(x)$. In general, an inference rule pre-existed its related schema. Thus, what one should infer depends on how one describes a given situation, and not on formal logic.

From this perspective, it is necessary to identify what determines a valid description of the situation. Next, we shift to the question of how descriptive statements arise.

Origin of Description

In mathematics, some descriptive statements are contained within the axioms of the system under consideration, but even in advanced mathematics, we do not always think in completely formalized systems. We propose that, instead, descriptive statements originate from *models* in our minds. In the present paper, the term *model* has a dual meaning. In this regard, Mason's (1989) idea is highly suggestive. According to Mason (1989), mathematical abstraction is described as "a delicate shift of attention from seeing an expression *as* an expression of generality, to seeing the expression *as* an object or property" (p. 2, italics in the original). Using the idea of "a shift of attention," we will show the dual meaning of "model."

One meaning is "something that a copy can be based on because it is an … example of its type" ("Model," n.d.-a). We call this an *exemplary model*. For example, the set of all integers, together with the operation $+$, is an exemplary model of a group in our minds, because it is a typical example of a group. With this in our minds, we can easily understand any example of a group by analogy. We can also show that the set of all integers with the operation $+$ is an exemplary model satisfying the group axioms. Similarly, because the experience of typicality can depend on subjective experiences, any example of a group can be an exemplary model. As it has not only the essential features of a group, but also non-essential ones, it has more information than a group as an abstract object without any non-essential features of a group. In general, an exemplary model satisfies a certain set of axioms, and carries more information than an abstract object without any properties which the axioms do not imply. A set of axioms do not have to be commonly accepted. Arbitrary logical expressions may be axioms. If a set of axioms is consistent, there exists at least one exemplary model for them.

Another meaning of the term "model" is "something that represents another thing … as a simple description that can be used in calculations" ("Model," n.d.-b). We call this a *descriptive model*. For example, a line in mathematics may be regarded as a descriptive model of a physical line, such as that made by a pencil, in our minds. A line in mathematics is defined by focusing attention on only some of the features of a physical line. It is a result of neglecting uninteresting features that. While a physical line does have width, we usually require in mathematics that a line have no width. In general, a descriptive model is created by focusing attention on only some of the features of other descriptive models or physical objects. Such a temporal creation is then refined with certain provisos (e.g., "it has no width"). The provisos prevent us from focusing attention on uninteresting features of the source descriptive models or objects.

Most relevant here is the relativity between exemplarity and descriptiveness. That is, when we focus attention on some essential features of an exemplary model, the abstract object constrained by the logical expressions of those features is a descriptive model of the exemplary model. When we create a new object by adding some extra features to an abstract object that is a descriptive model, the new object is an exemplary model of the descriptive model. In other words, any model in our minds can always be both exemplary and descriptive. Any model other than a physical object is an exemplary model of more abstract models or objects, and it is simultaneously a descriptive model of more concrete models or objects. The relativity between exemplarity and descriptiveness allows us to dispense with the distinction between the terms "model" and "object." In this sense, both terms may be used interchangeably, because every model can become an object of thought, and vice versa.

By using the term "model," one of the predominant origins of descriptive statements in mathematics can be explained as descriptions of models in our minds. We will provide two examples: the fundamental theorem of cyclic groups, and the construction of an equilateral triangle. Let us explain their possible models, for example, in the author's mind.

The fundamental theorem of cyclic groups: The theorem states that every subgroup of a cyclic group is cyclic. Let $\langle g \rangle$ be a cyclic group generated by g. Following the definition of a cyclic group, $\langle g \rangle$ simply consists of $\cdots, g^{-2}, g^{-1}, e, g, g^2, \cdots$; there is no other element in $\langle g \rangle$. If a subgroup of $\langle g \rangle$ has n different elements, they can be represented by $g^{k_1}, g^{k_2}, \cdots, g^{k_n}$. From the group axioms, the subgroup contains $g^{\mathrm{GCD}(k_1, k_2, \cdots, k_n)}$, and $g^{\mathrm{GCD}(k_1, k_2, \cdots, k_n)}$ generates all elements in the subgroup. Thus, the theorem seems to be true.

This way of creating descriptions of models in our minds implies various prescriptions. For example, when someone says that $\langle g \rangle$ might not contain e, the author should argue that $\langle g \rangle$ always contains e because $\langle g \rangle$ is an example of a group. As another example, when someone points out that the order of a subgroup of $\langle g \rangle$ is not always finite, the author should recognize that an example of a subgroup of $\langle g \rangle$ in his mind is too specific.

The construction of an equilateral triangle on a given line segment: Let AB be the given line segment. Draw a semicircle with center A and radius AB. Again, draw a semicircle with center B and radius BA on the same side as the first semicircle. Let C be the point of intersection of the semicircles. Then, the triangle ABC is equilateral. This is because the semicircles centered at A and B have radii of equal length, and all three segments AB, BC, and CA are the length of their radii. Thus, the construction seems to be valid.

There are also various prescriptions in this case. For example, when someone says that the three edges AB, BC, and CA are not always equal, the author should argue that they are always equal, for the following reason. The point C is regarded as our exemplary model of the points on the semicircles A and B; the pairs CA, AB and AB, BC are regarded as our exemplary models of equivalent radii, and the lengths of AB, BC, and CA are regarded as our exemplary models of the transitivity rule. As another example, if someone points out that the author's consideration depends on the belief that the two semicircles always intersect with each other, he should recognize that his consideration depends on a visual representation.

Generally speaking, descriptive statements of some mathematical objects are created by accessing their models in human minds, and then describing these models. Given an axiomatic system (that is, a descriptive model), one creates an exemplary model of the given descriptive model in mind. Creating a descriptive statement in the system is creating a descriptive model of the exemplary model. There are two types of creation. One creates a description of a common property among all the exemplary models of the given descriptive model. The other creates a description of a property satisfied by only a particular exemplary model of the given descriptive model. If one mistakenly argues something based on the latter type, and someone points this out, then one should recognize the mistake (for example, that an example of a subgroup of $\langle g \rangle$ is too specific, or the consideration of an equilateral triangle depends on a visual representation). Descriptive statements in mathematics, therefore, can come from descriptions of models in our minds, and prescriptive statements can come from the exemplarity of the exemplary models. From this perspective, the reason proofs and refutations (Lakatos, 1976) occur in mathematics might be because humans (including mathematicians) sometimes create a description of a property satisfied by only a particular exemplary model of the given descriptive model.

Conclusion

For the purpose of building a supplementary theory from which we can derive practical suggestions for fostering inquiring minds in mathematics, this paper proposed a theoretical perspective to explain the origin of the discoverability belief. The main results are as follows: (i) We can simultaneously treat descriptive statements and prescriptive statements. (ii) We can distinguish between descriptive and exemplary models. (iii) Descriptive statements in mathematics can come from the descriptions of models in our minds, and prescriptive statements can come from the exemplarity of exemplary models.

We now argue that the discoverability belief arises from prescriptiveness in mathematics, and that three implications can be derived from our proposed perspective and existing theoretical perspectives. First, even taking any standpoint criticized by Goldin (2003) (such as radical constructivism (e.g., von Glasersfeld, 1995), social constructivism (e.g., Ernest, 1998), or the perspective of embodied cognition (e.g., Lakoff & Núñez, 2000)), one can treat prescriptiveness in mathematics, because individual thinking creates its own prescription. Second, the roles of action and process in mathematics become more important. These roles have been emphasized in APOS theory (e.g., Dubinsky & McDonald, 2002) and in Tall's framework (e.g., Tall, 2008). From the proposed perspective, descriptive statements can be created through a shift of attention from the particularity to the essential features of a mathematical object, and the roles of action and process may be interpreted as the effect of defocusing from nonessential features. We tend to have an interest in some invariant properties of all the elements under consideration. Third, the proposed perspective is also compatible with the new materialist ontologies of de Freitas and Sinclair (2013), though we may need to reconsider what kinds of situations can transfer consistency from physical reality to mathematics. This is because an individual and subjective shift of attention does not always warrant the transference of consistency. In summary, each of the existing perspectives is compatible with the proposed one.

As a practical suggestion from the proposed perspective, we point out that students might lose the discoverability belief if they recognize the refutation of their argument as a denial of their way of mathematical thinking. What the refutation actually denies might not be their attitude toward creating an exemplary model of the given descriptive model, but only the particular exemplary model contingently created at that time. If creating an exemplary model and describing it is an essential process of mathematics, a chain of reasoning means a chain of creating exemplary models or descriptive models of the already-created models. Then, many chains of reasoning are not deductive. If a student seems to mistakenly make a non-deductive chain of reasoning, the teacher should carefully communicate with the student, and try to recognize which chain would make such a conclusion. Otherwise, proofs and refutations do not work well as a social construction of mathematical knowledge in classrooms, and intersubjectivity cannot be established. In particular, it seems to be important for the teacher to pay attention not only to the student's conclusion but also to their attitude toward developing new findings in order to foster inquiring minds in mathematics.

There are at least two limitations of the proposed perspective. First, it is still not clear whether it is completely compatible with each existing research perspective. Second, the above practical suggestion is still based on assumptions whose validity is not always warranted (for example, whether reasoning always means creating models). The suggestion describes only a possible situation in classrooms. Further development of our theoretical framework in this regard provides an avenue for future research.

Acknowledgment

This work was supported by Grant-in-Aid for JSPS Fellows No. 252024. Any opinion stated here is that of the author and does not necessarily reflect the views of JSPS.

References

De Freitas, E., & Sinclair, N. (2013). New materialist ontologies in mathematics education: the body in/of mathematics. *Educational Studies in Mathematics*, *83*(3), 453–470.

Dubinsky, E., & McDonald, M. A. (2002). APOS: A Constructivist Theory of Learning in Undergraduate Mathematics Education Research. In D. Holton, M. Artigue, U. Kirchgräber, J. Hillel, M. Niss, & A. Schoenfeld (Eds.), *The Teaching and Learning of Mathematics at University Level* (Vol. 7, pp. 275–282). Dordrecht: Kluwer Academic Publishers.

Ernest, P. (1998). *Social Constructivism as a Philosophy of Mathematics: New Essays in the Comparative Philosophy of Religions*. SUNY Press.

Fallis, D. (2003). Intentional gaps in mathematical proofs. *Synthese*, *134*(1), 45–69.

Goldin, G. A. (2003). Developing complex understandings: On the relation of mathematics education research to mathematics. *Educational Studies in Mathematics*, *54*(2-3), 171–202.

Lakatos, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press.

Lakoff, G., & Núñez, R. E. (2000). *Where Mathematics Comes from: How the Embodied Mind Brings Mathematics Into Being*. Basic Books.

Mason, J. (1989). Mathematical abstraction as the result of a delicate shift of attention. *For the learning of mathematics*, *9*(2), 2–8.

Model. (n.d.-a). In *Cambridge Dictionaries Online*. Cambridge University Press. Retrieved from http://dictionary.cambridge.org/dictionary/british/model_1

Model. (n.d.-b). In *Cambridge Dictionaries Online*. Cambridge University Press. Retrieved from http://dictionary.cambridge.org/dictionary/british/model_4

Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal*, *20*(2), 5–24.

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. The Flamer Press.

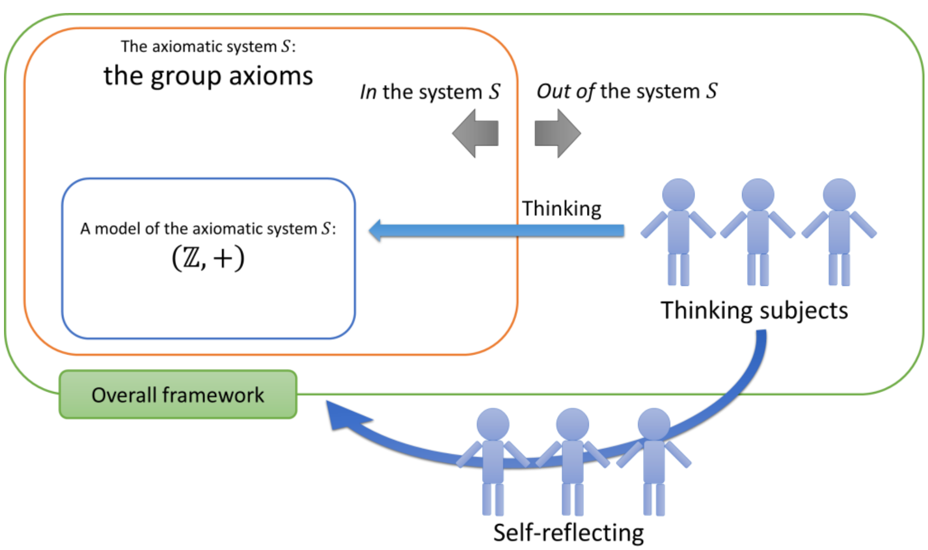


Figure 1. Relationships between the inside and the outside of a system

EXPLORING DIFFERENCES IN TEACHING PRACTICE WHEN TWO MATHEMATICS INSTRUCTORS ENACT THE SAME LESSON

Joseph F. Wagner
Xavier University

Karen Allen Keene
North Carolina State University

Investigating teacher practice at all educational levels has become an important research arena. We analyze teacher practice by comparing two implementations of the same fragment of a student-centered curriculum by two mathematics professors. We highlight differences in their practices and the consequent classroom results by analyzing their participation in class discussions, and we show how Schoenfeld's (2011) resources, goals, and orientations framework may be used to explain these differences. Using classroom and interview data, we identify resources that each instructor believed he lacked, we highlight prominent mathematical and social goals that each instructor held, and we infer orientations toward teaching and learning mathematics that guided each instructor's practices. All of these in combination suggest explanations for the observed differences in the implementations and class outcomes. We believe that this analysis provides an important technique to understand and improve teaching and learning at the undergraduate level in mathematics.

Key words: Teaching Practice, Teacher Beliefs, Teacher Goals, Student-Centered Instruction

Science, technology, engineering, and mathematics (STEM) disciplines have been increasingly identified as a priority for educational improvement and innovation in the United States. The Department of Commerce (2012) listed mathematics and science education as one of six alarms that require our utmost attention in the 21st century. One way to improve STEM education is to improve mathematics teaching practice at the university level. To do this, researchers need to develop a variety of perspectives from which to analyze what takes place in college classrooms.

Our work has focused on the following two research questions:

- How do university professors practice the art and science of teaching when engaging in a new (for them) innovative student-centered differential equations curriculum?
- What factors, particularly, their personal resources, goals and orientations, influence their instruction in the classroom?

There is considerable research on teaching practice at the elementary and secondary level. Ball and Forzani (2009) indicated that study of "the work of teaching" is a particularly important area. We believe that it is also important to study university teaching practice, especially as more innovative curricula are introduced to the college level. In this presentation, we report on the results of our investigation of two university mathematicians' implementation of an innovative, student-centered curriculum. Because of the magnitude of the available data, we have chosen to focus on only the first day of the implementation. We provide a detailed analytical snapshot of the mathematicians' practices, and we use Schoenfeld's (2011) resources, goals, and orientations framework to offer an explanation for the differences we observed.

Literature Review and Theoretical Framework

During the past few years, researchers have begun to respond to Speer, Smith, and Horvath's (2010) call for more empirical research in the *practice* of mathematics teaching at the university level. Some have examined mathematics professors' practice using lecture methods (see Trenholm, Alcock, & Robinson, 2012), and of increasing interest are studies of mathematicians changing their teaching practices from lectures to a more student-centered approach. Wagner, Speer, and Rossa (2007) reported on one instructor's knowledge as he implemented an inquiry oriented DE course. They identified forms of knowledge apart from mathematical content knowledge that are essential to reform-oriented teaching, and highlighted how knowledge acquired through more traditional instructional practices may fail to support research-based forms of student-centered teaching. Speer and Wagner (2009) expanded the study of the same mathematician's teaching by connecting pedagogical content knowledge and mathematical knowledge through the construct of analytic scaffolding. Finally, Wagner (2007) presented preliminary work at RUME about two mathematicians and their differences. Our work has refined and developed this earlier preliminary analysis.

Lee et al. (2009) suggested the construct of *mathematical content move* to discuss one mathematician's practice while first implementing a student-centered differential equations curriculum. They identified the instructor's mathematical agenda which may have influenced his practice and offered five specific moves that the teacher made. Johnson et al. (2013) discussed their case studies of three mathematicians. They found that three themes emerged from interviews and reflections with these mathematicians: curriculum coverage; goals for student learning; and the role of the teacher. Our current work distinguishes itself, however, by contrasting the practices of two teachers implementing identical curriculum materials.

Schoenfeld's (2011) model of goal-oriented decision making suggests that important aspects of teachers' practices can be understood as a function of their resources, goals, and orientations (such as beliefs or preferences). Teachers' orientations frame their perceptions, influence their goals and the prioritization of those goals, and activate relevant resources, particularly their knowledge. Decisions consistent with the goals are made, consciously or unconsciously, about which teaching directions they will focus on and which resources, such as their knowledge and skills, are most needed. Our immediate goal is not to explain each decision of a given teacher, but to identify salient aspects of their knowledge, goals, and orientations that may explain equally salient characteristics of the teacher's practice as witnessed throughout a classroom episode.

Several other researchers have recently attempted to use Schoenfeld's framework to analyze the practice of university teachers (e.g., Hannah, Stewart, & Thomas, 2011; Patterson, Thomas, & Taylor, 2013; Törner, Rolka, Rösken, & Sriraman, 2010), however, we are aware of only one other attempt to use this framework to compare the practices of two instructors using the same lesson materials. Pinto (2013) compared the lessons of two TAs who individually interpreted and implemented the same lesson plan very differently. Speer, Smith, and Horvath (2010) cited an explicit need for comparison studies of teacher practice at the university level, and we believe our data provide an ideal opportunity to contribute to this research need.

Methods

Data for the current study are taken from a much larger collection gathered as Prof. X and Prof. Y each taught a semester course in Differential Equations, two years apart, at a private, liberal arts university in the Midwest. The students in the class were primarily majors in Mathematics or one of the physical sciences. Both instructors had doctorates in Mathematics and each had been teaching for over 15 years at the university level. Both used the same set of curricular materials for an Inquiry-Oriented Differential Equations (IO-DE) course developed by Rasmussen (2006).

Almost all of their classes were videotaped with two cameras, one following the instructor and another focused on a selected small group of students. Audiotaped interviews were conducted with each instructor by the first author several times prior to the semester and after almost every class. For the present study, complete transcripts were made of the whole-class discussions for each instructor's first day of class, and significant portions of the interviews carried out near the first day of class were also transcribed.

The instructors' contributions to the whole-class conversations were coded using a coding scheme inspired by Wells and Arauz (2006) to determine the role that each turn of talk played in the conversation. The codes, designed to capture the nature of each comment and each question, will be described in the presentation. The two authors coded the transcripts independently using 18 possible codes, with 72% and 73% agreement for Prof. X's class and Prof. Y's class respectively. Disagreements were resolved by mutual discussion.

Both interview and classroom transcripts were searched for statements that offered implicit or explicit insight into important aspects of each instructor's knowledge, goals, and orientations. Particular attention was given to recurring themes or ideas as a form of triangulation of the evidence supporting our claims. We present our results with numerous and substantial excerpts from the transcripts.

Analysis

The first part of our analysis summarizes some key aspects of what took place during the whole-class discussions in each instructor's class. The second part (limited for this proposal) is an example of the analysis of each instructor's resources, goals, and orientations that we believe contributed to the differences in the two classes. Our complete analyses are considerably more comprehensive.

An overview of the two classes

From a broad perspective, the two classes looked a great deal alike, with both instructors demonstrating clear attempts to use the IO-DE materials with fidelity to the developers' intent. Both classes used identical problems provided by the curriculum for the first day of class. Despite the overall similarities, however, the instructors participated in the class discussions very differently, and the two classes played out with different outcomes. Our coding of each question and comment made by each instructor during whole-class discussions is summarized in Table 1. (Details on the coding categories will be provided during the presentation.) We also analyzed the rate and breadth of participation in the conversations by the instructors and the students. The times spent in the two classes on whole-class discussions were nearly identical for both instructors, so quantitative comparisons were made easily.

In short, we found that Prof. Y's class was marked by widespread student participation in whole-discussions, with a significant majority of the students making at least one contribution, and with no students tending to dominate the conversation. Prof Y's questioning style was highly open-ended and non-directive, intentionally inviting students to offer their ideas, without giving feedback concerning the correctness or incorrectness of those ideas. He rarely asked mathematically pointed questions, and he rarely focused on something a student said except to invite very general feedback from other students. The content of the conversation was consequently wide and varied, and little to no discernible direction was evident, resulting in an absence of any agreed-upon conclusion to the first (and only) problem discussed, even at the conclusion of the class.

Prof. X's class was also animated by student participation, but proportionally fewer students contributed to the whole-class discussions, and most of the contributions were made by the five most vocal students. Prof. X's questioning style included open-ended questions inviting students to share their ideas, but it was also marked by a large number of questions that were mathematically pointed. Prof. X often took a student's comment and constructed a more specific follow-up mathematical question from it, thereby focusing and directing the conversation. He used evaluative feedback, occasionally shared his own ideas, and sometimes gave answers to questions that students did not answer themselves. By the end of the class, Prof. X had indicated that the class appeared to have agreed on the answers to the first two problems of the curricular sequence, and a third problem was discussed. When no consensus concerning the third problem was evident, Prof. X announced and briefly explained the correct answer just as the class ended.

	Prof. Y	**Prof. X**		**Prof. Y**	**Prof. X**
Questions			**Comments**		
Thinking	32	19	Direction	12	10
Opening	24	10	Observation	16	13
Neutral	0	1	Revoice	7	8
Narrowing	8	8	Summarize	4	1
Math/Service	2	26	Evaluate	0	7
Clarify	12	4	Opinion/Thinking	0	2
Progress/Assess	13	8	Tell	0	2
Justify	0	6	Joke	0	12
Corrective	1	1	Other	5	6
Other	6	2			

Table 1: Coding counts for instructors' questions and comments

Explaining the differences

To explain the differences we observed in the instructors' practices and the subsequent class outcomes, we turn to an analysis of each instructor's resources, goals, and orientations, as gleaned from the instructors' own words transcribed from interviews and class discussions. In our complete analysis, we offer evidence of significant differences in what the two instructors perceived to be lacking in their own knowledge resources that affected their behavior on the first day of class. In addition, we consider evidence of their contrasting goals (both social and mathematical) and orientations (particularly beliefs). Our objective is to use these analytical lenses to construct a compelling explanation for how each instructor's constellation of resources, goals, and orientations functioned together to result in some of their key classroom practices and, hence, the different outcomes of the classes and different opportunities for student learning.

Due to limitations of space, we provide here only a condensed example of how one notable difference in the two class outcomes can be explained in this manner. The coding tallies in Table 1 for "Evaluate," "Opinion/Thinking," and "Tell" show that not once during his entire class did Prof. Y make any comment that evaluated the mathematical correctness of a student's contribution, offer his own opinion about the answer to a mathematical question or describe his own thinking about the mathematical ideas being discussed, or tell the students the correct answer to a question or problem that had not been answered correctly by them. Prof. X, on the other hand, made seven evaluative responses to students, twice described his own way of thinking about how a problem could be solved, and twice told the students answers to a mathematical question that they had not answered on their own. We examine these three codes together because they all relate to how each instructor exercised his own mathematical authority in the classroom.

Prof. Y indicated throughout his interviews that he believed that if students are to learn mathematics with real understanding, then it is important for them to construct their knowledge on their own. In fact, it was primarily for this reason that he wanted to use the IO-DE curriculum materials, since he believed that the student-centered activities could promote this kind of understanding. When discussing his goals for the first class, he was clear on his desire to focus on social goals, primarily by immediately getting norms in place that encouraged students to share their ideas aloud and with each other, and to respond and critique each other's ideas, appropriately. When asked about his mathematical content goals for the class, he had a difficult time naming any:

> I was kind of passing things around without having too clear of a goal … But really what I wanted to do was discuss it. And exactly what comes up is not maybe that important.

At the same time, Prof. Y had strong beliefs about his role in the class discussions—particularly that he should not exercise his mathematical authority in any way, lest the students become dependent on him rather than on their own good reasoning.

> I didn't want to, you know, start telling them in any way, not in any way. […] I didn't really want to represent the truth.

Prof. X also exhibited a conviction that students needed to construct their own mathematical understanding, so he, too, had goals to establish certain social norms on the first day. However, he also showed evidence of having articulated for himself some clear mathematical goals for the class. For example, when asked if he had planned on a particularly insightful idea arising from his students, he replied "I hoped that it would. It was certainly in the notes that I wrote for myself." Like Prof. Y, Prof. X was also concerned about not exercising mathematical authority:

> What I don't want to do is to, I don't want to lay down the truth at some point, because then they'll say, "We'll stumble around to wherever we get, but then we'll always depend on him to lay down the truth before we move on."

Unlike Prof. Y, however, Prof. X greatly tempered that restriction because of his clear and strong belief that learning mathematics with understanding takes time and struggle, and that it would not be reasonable for him to expect students to understand a lot in the classroom:

> I don't know that it's important to me that everybody, or even a significant proportion of people, get it at the moment. I think it's their responsibility to get it in a variety of ways. It comes from my own personal experience with mathematics.

As a result, Prof. X, with his clear mathematical goals for the class in mind, was more inclined to direct the mathematical ideas forward, even if it meant exercising greater authority. Students *did* indeed need to construct their own understandings, but he believed that most of this would have to take place outside of class. Prof. Y, on the other hand, with his primarily social goals at the fore, was much less concerned about how far along in the curriculum he got (at least on the first day), and more concerned that students share and develop their own ideas without any exercise of authority on his part.

Conclusion and Implications for Research

Our research shows that Schoenfeld's framework for understanding why teachers do what they do offers a good way to explain the differences in two enactments of one day of a student-centered differential equations curriculum. In this proposal, we illustrated how differences in the instructors' goals and orientations effectively explain one divergence in their classroom practice. Because one of the instructor's primary goals was social, to create an environment where students' constructed mathematics by participation, and the other one's belief was that most students' do not learn mathematics for mastery the first time they see it, the instructors exercised their mathematical authority in different ways, and the learning opportunities and outcomes of the two classes were significantly different.

Research has shown that instructor practice is important to study in order to understand and improve student learning (Ball & Franzani, 2009). However, it is not clear that what is known to be the case at the K-12 level carries over to the university level. For example, despite the growing influence of student-centered activities, we have no clear answers to a question faced by these instructors: How much understanding of conceptually challenging university-level mathematics can we expect students to understand and develop with the time frame of an inquiry activity? Further, although we know that teachers' beliefs and attitudes significantly affect the enactment of a curriculum at the K-12 level (e.g. Collopy, 2003; Stipek, Givvin, Salmon, & MacGyvers, 2001; Arbaugh, Lannin, Jones, & Park-Rogers, 2006), even the term *curriculum* is used less often at a university level, where *content* is primarily the teacher's focus. We believe that as researchers are still only beginning to study teacher practice at the university level, issues such as these must be given greater attention.

References

Arbaugh, F., Lannin, J., Jones, D., & Park-Rogers, M. (2006). Examining instructional practices in Core-Plus lessons: Implications for professional development. *Journal of Mathematics Teacher Education*, *9*(6), 517–550.

Ball, D. & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education,* 60. p. 497-511.

Collopy, R. (2003). Curriculum materials as a professional development tool: How a mathematics textbook affected two teachers' learning. *Elementary School Journal*, *103*(3), 287–311.

Department of Commerce (2012). The competitiveness and innovative capacity of the United States. Report retrieved 08/08/2013 at http://www.commerce.gov/americacompetes.

Hannah, J., Stewart, S. & Thomas, M. (2011): Analysing lecturer practice: the role of orientations and goals, *International Journal of Mathematical Education in Science and Technology,* 42:7, 975-984.

Johnson, E. Caughman, J. Fredericks,, J. & Gibson, L. (2013). Implementing inquiry-oriented curriculum: From the mathematicians' perspective. *Journal of Mathematical Behavior.*

Lee, H. S., Keene, K. A., Lee, J. T., Holstein, K., Early, M.E., Eley, P. (2009). Pedagogical content moves in an inquiry-oriented differential equations class: Purposeful decisions to further mathematical discourse. In Proceedings of the *Annual Conference on Research in Undergraduate Mathematics Education conference*, Raleigh, NC.

Patterson , J.,Thomas, M. & Taylor, S. (2011): Decisions, decisions, decisions: what determines the path taken in lectures? *International Journal of Mathematical Education in Science and Technology,* 42:7, 985-995.

Pinto, A. (2013). Revisiting university mathematics teaching: A tale of two instructors. Presented at the Eighth Congress of European Research in Mathematics Education (CERME 8), Antalya, Turkey.

Rasmussen (2006). Inquiry-Oriented Differential Equations, Unpublished instructional materials.

Schoenfeld, A. H. (2011). *How we think: A theory of goal-oriented decision making and its educational applications*. New York: Routledge.

Speer, N. M., Smith, J. P., III, & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, 29, 99–114.

Speer, N. M. & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. *Journal for Research in Mathematics Education 40*(5), 530-562.

Stein, M., Remillard, J., & Smith, M. (2007). How curriculum influences student learning. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 1, pp. 319–369). Charlotte, NC: Information Age Publishing.

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, *17*(2), 213–226.

Törner, G., Rolka, K., Rösken, B., & Sriraman, B. (2010). Understanding a teacher's actions in the classroom by applying Schoenfeld's theory *Teaching-in-Context*: Reflecting on goals and beliefs. In B. Sriraman & L. English (Eds.), *Theories of mathematics education: Seeking new frontiers* (pp.401–420), Berlin/Heidelberg: Springer.

Trenholm, S., Alcock,L. & Robinson, C. (2012) Mathematics lecturing in the digital age, *International Journal of Mathematical Education in Science and Technology*, 43:6, 703-716.

Wagner, J. F. (2007, February). A comparison of two mathematicians' use of an inquiry-oriented differential equations curriculum. Presented at the Tenth Conference on Research in Undergraduate Mathematics Education (RUME), San Diego, CA.

Wagner, J. F., Speer, N. M. & Rossa, B. (2007). Beyond mathematical content knowledge: A mathematician's knowledge needed for teaching an inquiry-oriented differential equations course. *Journal of Mathematical Behavior*, *26*, 247-266.

Wells, G. & Arauz, R. M. (2006). Dialogue in the classroom. *Journal of Learning Sciences*. 15. pp. 379-428.

RAISING CALCULUS TO THE SURFACE: DISCOVERING MULTIVARIABLE CALCULUS CONCEPTS USING PHYSICAL MANIPULATIVES

Brian Fisher
Pepperdine University

Jason Samuels
Borough of Manhattan Community College

Aaron Wangberg
Winona State University

Eric Weber
Oregon State University

Current research on algebraic and quantitative reasoning shows that many students experience mathematics as the manipulation of meaningless symbols (Smith & Thompson, 2007). In order to develop meaning in symbolic contexts, students must first conceive of relationships between the underlying quantities present in a particular context. Our project focuses on a quantitative reasoning approach to multivariable calculus, in particular the concepts of function, rate, area and volume by using physical surfaces. In this poster, we provide examples of identifying, measuring, and recording of essential quantities on physical surfaces.

Key words: [multivariable calculus, quantitative reasoning, rate of change, multiple representations, physical model]

Current research on algebraic and quantitative reasoning shows that many students experience mathematics as the manipulation of meaningless symbols (Smith & Thompson, 2007). In order to develop meaning in symbolic contexts, students must first conceive of relationships between the underlying quantities present in a particular context. Our project focuses on a quantitative reasoning approach to multivariable calculus, in particular the concepts of function, rate, area and volume by using physical surfaces. In this poster, we provide examples of identifying, measuring, and recording of essential quantities on physical surfaces.

Multivariable calculus introduces significant complexity in quantitative reasoning, stemming from the need to think about interactions among many quantities and representing the variation in those quantities using variables. In response to this added complexity, Weber (2012) found some students, coined as novice shape thinkers, failed to reason quantitatively about graphs of multivariable functions, instead arguing primarily in terms of the topographical features of the surface.

There is mounting evidence that students have difficulties with the three-dimensional visualization and the geometric thinking necessary for success in the sciences (e.g. NAP, 2006). Many students have difficulty reasoning spatially about two dimensional images which are supposed to represent three dimensional objects (Price & Lee, 2010). This difficulty includes 2D representations on paper or a computer screen (Dede et al. 1999), which can be technically challenging for students (Hubona et al. 1999).

The project presented in the poster utilizes six different models, or surfaces, which represent multivariable functions. Each surface, a part of which is shown in Figure 1, has a transparent top onto which students can draw with dry-erase markers. Mats depicting coordinate systems or the surface's contour lines, like those shown in Figure 2, can be placed underneath the surface, helping students transfer drawings between the surface and underlying mat. The surfaces and accompanying tools, like the inclinometer shown in Figure 1, help students measure change (e.g. gradient, partial derivatives, directional derivatives) and accumulation (e.g. line integrals, surface integrals). Accompanying the surfaces are activities designed to encourage student exploration of the key ideas in multivariable calculus.

Preliminary data suggests these models help develop students' abilities to reason quantitatively and spatially while moving between multiple representations of multivariable functions. The project's manipulatives are designed to encourage student exploration between the different representations of multivariable functions and to connect quantitative features to algebraic representations. In addition, students are able to measure rates of change on the surface and connect their understanding of a single-variable derivative to the various derivatives introduced in multivariable calculus

During our poster presentation we will share samples of the physical manipulatives (surfaces, mats, and measurement tools) and activities designed to help students identify important mathematical concepts introduced in multivariable calculus. Participants will be invited to draw on the surfaces, try the activities, and explore concepts with the surfaces. In addition to informing conference attendees of our project, we would like to solicit feedback from the community on our project and ways in which we could incorporate these models into the mathematics and science curriculum.

References

Dede C., Salzman M., Loftin R., Sprague D. (1999). Multisensory immersion as a modeling environment for learning complex scientific concepts. In: Roberts N, Fuerzig W (eds) *Computer modeling and simulation in science engineering*. Springer, New York, pp 282–319.

Hubona G., Wheeler P., Shirah G., Brandt M. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer-Human Interaction,* v6, p214–242.

NAP - Committee on Support for Thinking Spatially. (2006). Learning to think spatially. Washington, DC: National Academies Press.

Price, Aaron & Lee, Hee-Sun (2010) The Effect of Two-Dimensional and Stereoscopic Presentation on Middle School Students' Performance of Spatial Cognition Tasks. *Journal of Science Education and Technology*, v19 n1 p90-103.

Smith, J., & Thompson, P. (2007). Quantitative reasoning and the development of algebraic reasoning. In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the early grades* (pp. 95-132). New York: Erlbaum.

Weber, E. (2012), "Students' Ways of Thinking about Two-Variable Functions and Rate of Change in Space" (Doctoral Thesis). Retrieved from http://pat-thompson.net/PDFversions/Theses/2012Weber.pdf on March 28, 2012.

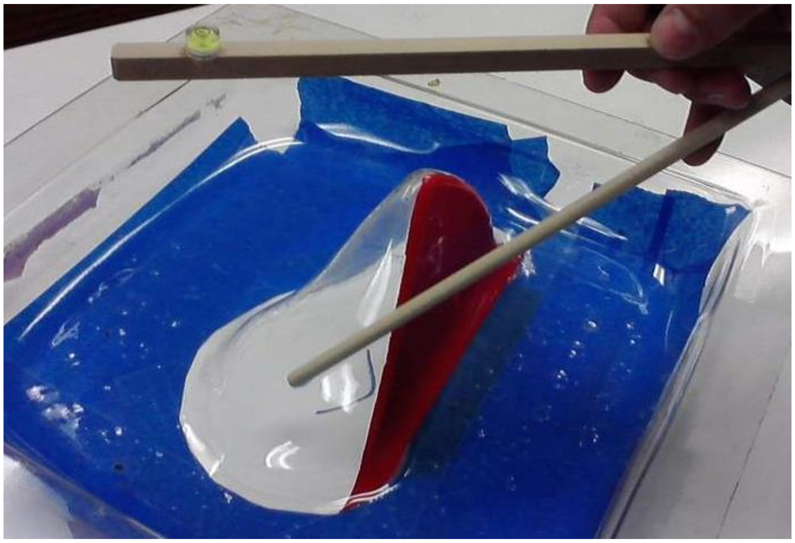


Figure 1: Measuring rate of change on the surface along a curve. (The picture depicts part of the new transparent surface models.)

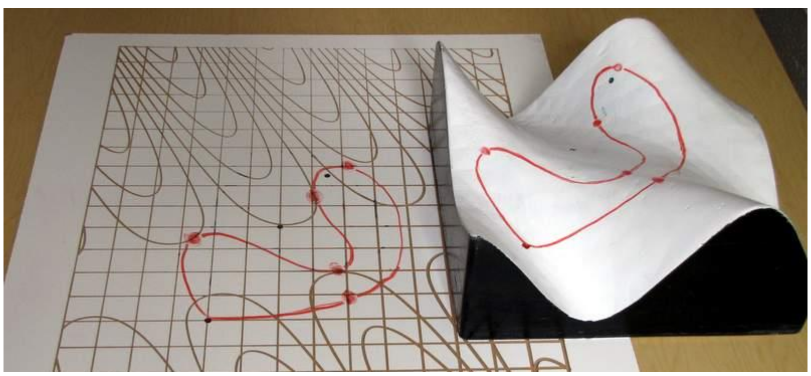


Figure 2: Model surface (original, non-transparent model) and contour grid.

DEVELOPING INQUIRY ORIENTED INSTRUCTIONAL MATERIALS FOR LINEAR ALGEBRA (DIOIMLA): OVERVIEW OF THE RESEARCH PROJECT

Megan Wawro
Virginia Tech

Michelle Zandieh
Arizona State University

Chris Rasmussen
San Diego State University

Christine Larson
Florida State University

David Plaxco
Virginia Tech

Katherine Czeranko
Arizona State University

The goals of the recently funded DIOIMLA research project are to produce: (a) student materials composed of challenging and coherent task sequences that facilitate an inquiry-oriented approach to the teaching and learning of linear algebra; (b) instructional support materials for implementing the student materials; and (c) a prototype assessment instrument to measure student understanding of key linear algebra concepts. Our poster will provide more detailed information about the DIOIMLA research project. Each of the three aspects of the project will be described in more detail and examples of each will be shared. The poster will also include an overview of the current status of the research project and a summary of the timeline for planned future work.

Keywords: Linear algebra; curriculum design; inquiry oriented instructional materials; assessment

Linear algebra is widely viewed as pivotal yet difficult for university students, and hence innovative instructional materials are essential. The goals of the recently funded DIOIMLA research project are to produce: (a) student materials composed of challenging and coherent task sequences that facilitate an inquiry-oriented approach to the teaching and learning of linear algebra; (b) instructional support materials for implementing the student materials; and (c) a prototype assessment instrument to measure student understanding of key linear algebra concepts. The project makes a needed contribution to the field by developing instructional materials that allow for active student engagement in the guided reinvention of key mathematical ideas. It also develops instructional support materials that convey the instructional designers' intention without being overly prescriptive and that provide information about how students think and learn within the task sequences. The production of a prototype assessment instrument is of value because it furthers what is known about student thinking in linear algebra and provides a measure of comparison across pedagogical approaches. The study partners mathematics education researchers and mathematicians to incorporate research on teaching and learning into effective pedagogical approaches at the undergraduate level.

Prior Work and Theoretical Framing

The DIOIMLA research program builds from a previously NSF-funded project focused on student learning of basic ideas in linear algebra as students transitioned from intuitive to more formal ways of reasoning. Through conducting interviews and watching classroom video data, we analyzed and reported extensively on student thinking about particular mathematical ideas (e.g., Larson, Zandieh, & Rasmussen, 2008; Larson & Zandieh, 2013; Wawro, Larson, Zandieh, & Rasmussen, 2012; Wawro, Rasmussen, Zandieh, & Larson, 2013; Wawro, Rasmussen, Zandieh, Larson, & Sweeney, 2012; Wawro, Sweeney, & Rabin, 2011; Zandieh, Ellis, &

Rasmussen, 2012). This design research consisted of a cyclical process of ongoing analysis of student reasoning and simultaneous task design and conjecture modification regarding the possible paths that students' learning might take (Gravemeijer, 1994; Cobb, 2000). This prior work in linear algebra provides a strong foundation for all three goals of the DIOIMLA project.

Our theoretical framework for designing instructional materials draws on heuristics of Realistic Mathematics Education (summarized by Cobb, 2011). First, a task sequence should be based on experientially real starting points. Second, the task sequence should be designed to support students in making progress toward a set of associated mathematical learning goals. Third, classroom activity should be structured so as to support students in developing models-of their mathematical activity that can then be used as models-for subsequent mathematical activity. Finally, with instructor guidance, students' activity evolves toward the reinvention of formal notions and ways of reasoning about the mathematics initially investigated.

Purpose of the Poster

Our poster will provide more detailed information about the DIOIMLA research project. Each of the three aspects of the project –inquiry-oriented linear algebra student materials, accompanying instructor support materials, and a prototype assessment instrument – will be described in more detail and examples of each will be shared. The poster will also include an overview of the current status of the research project and a summary of the timeline for planned future work. The project team welcomes feedback from interested parties in the RUME community regarding any of the DIOIMLA research project.

References

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 307–330). Mahwah, NJ: Lawrence Erlbaum Associates.

Cobb. P. (2011). Learning from distributed theories of intelligence. In E. Yackel, K. Gravemeijer, & A. Sfard (Eds.), *A journey into mathematics education research: Insights from the work of Paul Cobb* (pp. 85-105). New York: Springer.

Gravemeijer, K. (1994). Educational development and developmental research. *Journal for Research in Mathematics Education, 25(*5), 443-471.

Larson, C., & Zandieh, M. (2013). Three Interpretations of the Matrix Equation *Ax=b. For the Learning of Mathematics, 33*(2), 11-17.

Larson, C., Zandieh, M., & Rasmussen, C. (2008, February). A trip through eigen-land: Where most roads lead to the direction associated with the largest eigenvalue. Paper presented at the 11th Conference on Research in Undergraduate Mathematics Education, San Diego, CA.

Wawro, M., Larson, C., Zandieh, M., & Rasmussen, C. (2012, February). A hypothetical learning trajectory for conceptualizing matrices as linear transformations. Paper presented at the 15th Conference on Research in Undergraduate Mathematics Education, Portland, OR.

Wawro, M., Rasmussen, C., Zandieh, M., & Larson, C. (2013). Design research within undergraduate mathematics education: An example from introductory linear algebra. In T. Plomp, & N. Nieveen (Eds.), *Educational design research – Part B: Illustrative cases* (pp. 905-925). Enschede, the Netherlands: SLO.

Wawro, M., Rasmussen, C., Zandieh, M., Larson, C., & Sweeney, G. (2012). An inquiry-oriented approach to span and linear independence: The case of the magic carpet ride sequence. *PRIMUS, 22*(8), 577-599. DOI: 10.1080/10511970.2012.667516

Wawro, M., Sweeney, G., & Rabin, J. M. (2011). Subspace in linear algebra: Investigating students' concept images and interactions with the formal definition. *Educational Studies in Mathematics, 78*, 1-19.
Zandieh, M., Ellis, J., & Rasmussen, C. (2012, February). Student concept images of function and linear transformation. Paper presented at the 15th Conference on Research in Undergraduate Mathematics Education, Portland, OR.

STUDENTS' USE OF PARAMETERS AND VARIABLES TO REASON ABOUT THE BEHAVIOR OF MULTIVARIABLE FUNCTIONS

Eric Weber
Oregon State University

The purpose of this paper is to characterize students' ways of thinking about parameters and variables to reason about the behavior of multivariable functions. I focus on two single variable calculus students, Lisa and Carl, as they participated in a sequence of semi-structured exploratory teaching interviews intended to gain insight into 1) their approaches to reasoning about the behavior of single variable functions, and 2) what role those approaches played in their initial thinking about the behavior of functions of two, three and four variables. The interviews suggest that the students' ability to move flexibly between thinking about a function's variables as parameters allowed them to generalize their reasoning patterns about functions of n variables and extend that to functions of n+1 variables. I argue that their ability to parameterize functions allowed them to reason about functions for which they could not initially visualize representations.

Keywords: Function, Representations, Graph, Quantitative reasoning, Way of thinking, Way of understanding.

Introduction

Physicists, chemists, engineers and biologists conceive of and use functions in unique ways, yet at the center of each of their uses is the notion that a function represents a relationship between quantities. In these disciplines, it is uncommon to have a function of one, or even two variables. For instance, physicists studying thermodynamics have so many variables that they must parameterize a number of a function's variables to reason about it. Engineers study complicated systems that are built on the assumption that hundreds of variables might be relevant. Similarly, students in these disciplines are required to reason about these complicated systems of relationships almost at the beginning of their programs. Yet much of what we understand about how students reason about functions focuses on simple systems of one variable. While students' reasoning about these systems is fascinating and provides novel insight into how to support their learning, it is not clear how their reasoning patterns extend to functions of more than one variable, and particularly to functions that cannot be easily represented by using a graph. Without understanding how students might come to develop the ways of reasoning we intend about multivariable functions, it is difficult to imagine instruction around these ideas changing. This particular study focused on students' reasoning about the behavior of multivariable functions and in what ways that reasoning depended on the behavior of single variable functions.

The purpose of this paper, and consistent with the study's aim, is to characterize students' ways of thinking about parameters and variables to reason about the behavior of multivariable functions. I focus on two first semester students, Lisa and Carl, as they participated in a sequence of semi-structured interviews intended to gain insight into 1) their approaches to reasoning about the behavior of single variable functions, and 2) what role those approaches played in their initial thinking about the behavior of functions of two, three and four variables. The interviews suggest that the students' ability to move flexibly between thinking about a function's variables as parameters allowed them to generalize their reasoning patterns about functions of *n* variables and extend that to functions of *n+1* variables. I argue that their ability to parameterize functions allowed them to reason about functions for which they could not initially visualize representations. Lastly, I consider the instructional implications of this work with a particular focus on functions that cannot be represented visually.

Theoretical Framework

Given the foci of the research questions on students' mathematical knowledge, I focus on Harel's (Harel, 2008a, 2008b, 2008c) characterization of mathematical knowledge as comprised of an understanding of mathematical content and mathematical practices, where the reflexivity between the two drives the development of each. This distinction comes from Harel (2008a, 2008b, 2008c), who proposed the DNR based instruction framework as a way to think about the learning and teaching of mathematics. The *duality principle* states that mathematical knowledge consists of both students' understanding of particular content in mathematics and the characteristics of their ways of thinking that influence their practice of doing mathematics. To clarify this distinction, Harel (2008a) articulated the notion of a *mental act*, which includes activities like interpreting, conjecturing, explaining, searching, and problem solving (p. 3). These mental acts are at the core of Harel's fundamental distinction between *ways of understanding* and *ways of thinking.* He proposed that a *way of understanding* is "a particular cognitive product of a mental act carried out by an individual" (p. 4) (Figure 1). He described a *way of thinking* as "a cognitive characteristic of a person's ways of understanding associated with a particular mental act" (Harel, 2008a). Harel's analogy was that ways of understanding correspond to subject matter knowledge and ways of thinking correspond to conceptual tools.

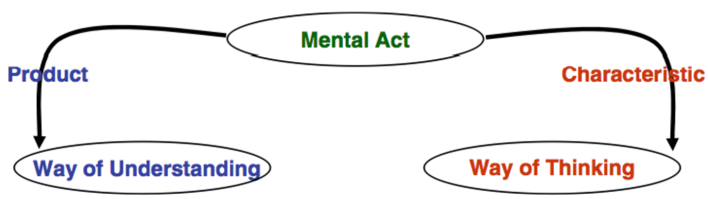


Figure 1. The mental act, way of understanding and way of thinking (Harel, 2008a)

A foundation of Harel's model of mathematical knowledge, illustrated in the duality principle, is that ways of thinking and ways of understanding are reflexive in nature. That is, "students develop ways of thinking only through the construction of ways of understanding, and the ways of understanding they produce are determined by the ways of thinking they possess" (Harel, 2008a). Ways of thinking, applied in one moment to a novel situation, can be thought of as a way of understanding. At the same time, repeated instances of ways of understanding might formulate patterns that develop into ways of thinking. This feedback between ways of thinking and ways of understanding is the core of the *duality principle*. Thus, this paper represents the development of students' mathematical knowledge in terms of ways of thinking, ways of understanding and reflexivity between the two.

Background Literature

Though the process by which students learn about two variable functions has not been fully explored, some researchers have characterized the representations students construct as they reason about tasks involving functions of two variables. For example, Yerushalmy (1997)argued that it is important to understand *what* is being generalized as students move from one to two variable functions. She studied six seventh grade algebra students in the context of describing, solving and generalizing about functions. She identified three key parts of the students' understanding of two variable functions: Identification of quantities and variation, generalizing graphical representation of a function in three dimensions, and manipulation of the algebraic conventions for defining a function. Yerushalmy found it essential that the students constructed the quantities under consideration, identified invariant relationships between the quantities, and imagined representing the invariant relationship using a function in both the one and two variable case.

As another example, Trigeruos & Martinez-Planell (2010, 2012) proposed and then refined a genetic decomposition of the understandings a student needs to conceive of two-variable functions and their graphs. They characterized students' notions of subsets of three-dimensional space and constructed an instrument and interview protocol to characterize what nine students knew about functions after taking a multivariable calculus course. Subsequently, Martinez-Planell and Trigueros (2012) proposed a genetic decomposition to postulate about constructions they students might make as they think about two variable functions. This genetic composition focused in depth on students' notion of planes in two and three dimensions as a basis for supporting students' notions of function in higher dimensions.

Both of these studies identified important understandings students need to think about graphs and representations of two-variable functions. Indeed, much of their data suggests that those representations were dependent on the ways in which the students conceived and then generalized their notion of single-variable functions behavior. There are two natural ways in which their foundational work might be extended. First, their results and conclusions suggest that generalization played a significant role in the students' conceptions of function behavior in three or more dimensions but because of their research questions generalization was not the central focus of their work. Second, they did not distinguish between ways of thinking and ways of understanding, the first of which appeared to be more powerful for students' ability to make generalizations than the second. This study focused explicitly on the role of ways of thinking and ways of understanding in the students' generalizations about function behavior from two to higher dimensions.

Method

The results described in this paper emerged from work with two single variable calculus students, Lisa and Carl, as they participated in a sequence of semi-structured exploratory interviews. These the purpose of these interviews was to gain insight into 1) their approaches to reasoning about the behavior of single variable functions, and 2) what role those approaches played in their initial thinking about the behavior of functions of two, three and four variables. Lisa and Carl were part of a larger group of six interviewees who volunteered to participate in the study, and I focus on them because of the stark difference in their approaches to reasoning about function behavior and how those ways of reasoning affected the generalizations they made about function behavior in higher dimensions.

It is important to distinguish between the interviews I completed and normal models of a teaching experiment in mathematics education. Typical teaching experiments continually generate and revise hypotheses about students' mathematical knowledge and use teaching as a means to both test and revise that model of their knowledge. The interviews I conducted for this study also included the generating and testing of hypotheses, but my intention was not to advance a particular way of reasoning for the students. Instead, I was concerned with characterizing the students' understanding of function behavior and describing how that affected their perception of function behavior in higher dimensions. Thus, while there were natural points at which I could have intervened instructionally, I did not do so.

During and after the interviews with six students, I used constant comparative analysis to identify emergent themes in the students' responses. The focus of the analysis was to identify students' ways of thinking and ways of understanding function behavior in two dimensions, and to understand in what ways those ways of thinking and ways of understanding affected their approaches to reasoning about function behavior in higher dimensions. Thus, the first coding scheme that emerged consisted of both ways of thinking and ways of understanding single variable functions. The second set of codes focused on students' ways of thinking and understanding of functions in higher dimensions. A comparison of the two coding schemes provided insight into the students' use of a way of thinking or way of understanding to reason about the behavior of functions. This comparison became the basis for identifying how ways of thinking and ways of understanding allowed students to generalize their notions of function behavior beyond two and three dimensions.

Results and Discussion

The predominant way of thinking that emerged from analyses of the interviews was the students conceiving of a function as an invariant relationship between variables. This way of thinking allowed the students to move flexibly between thinking about a variable as a parameter, and then a parameter as a variable when necessary. Their ability to hold variables constants (as parameters), allowed them to reason about the behavior of functions of many variables by making the function structurally similar to well known one-variable functions. However, the students' ability to see parameterization as necessary did not arise until they realized they could not described a multivariable function in the same way they could a one variable function. For instance, consider Lisa and Carl's responses to a task in which they were asked to describe the behavior of the function *f*, defined as $f(x_1,x_2)=x_1e^{x_2}$.

Carl: Well first of all this is really different for me. Hmm. I see two things. I see two functions I know, a linear one with just the first part, and then I see an exponential function with the second part. The difficulty I have initially is thinking about how to describe behavior because it cannot be both exponential and linear.

Lisa: The first thing I noticed here is that I know each piece of the function and can picture a graph and description at the same time. But I do not know how to coordinate the two pieces. For instance, I cannot say the rate of change is proportional to the amount of change like I can with exponential, because it is only true for half of the function.

This response is indicative of others given in interview one, in which both Carl and Lisa tried to describe the function in the way they would a one-variable function. Often they desired to use a single statement like "The rate of change of the function is 2", or "The rate of change is increasing at an increasing rate", which they found impossible for functions of two or more variables. During the remainder of interviews one and two, they began to focus less on describing function behavior with a single statement and more on describing pieces of the function. Their approach to describing a function in pieces drove their use of parameterization. As an example, consider their description the behavior of the following function f, defined as $f(x_1, x_2, x_3) = x_1 x_2^2 \cos(x_1 x_3)$.

Carl: Well, wow. So many things going on, I am almost forced to pick out things I know. So what I just treat everything but x-one as if is a number, or a given. Then the function becomes something that I know or can think about, a linear term and a cosine function that I can figure out. I could also do this for x-two, x-three, and get three separate types of functions. I don't really feel the need to combine all of those functions, as that just creates more of a mess as we have seen.

Lisa: Starting off, I see a bunch of things I recognize. If I was able to cover-up a few of the pieces of the function, almost like make them a given and focus on the rest, I could find three different functions in here. As I saw earlier though, I can't really put all three of those functions together, but I could give you descriptions of the behavior from them individually.

Lisa and Carl's ability to conceive of any variable as a parameter allowed them to reason about complicated functions of almost any number of variables. This was particularly useful for functions that could not be represented in two or three space. Indeed, Lisa and Carl both noted their previous dependence on graphical representations.

Carl: In the last interview we talked about functions of let's say 99 or 100, or *n* variables. Obviously we cannot picture these in our head; there is just no way to do it. So however many variables we have, I can give 1 less description, [e.g. 100 variables, 99] of the function's behavior. It makes me think how useless a graph can be sometimes!

Lisa: I think the most useful thing I have come up with in all of the interviews is that a graph is only possible for very few functions. To graph a function of 100 variables, you would have to hold constant 98 of the other variables! You would get so many graphs and have to coordinate them that it doesn't make sense any longer. I think my initial focus on describing a function in one statement just doesn't apply for these situations.

The most important way of thinking that emerged from analyses of the interviews was the students conceiving of a function as an invariant relationship between variables. Knowing that the function was an invariant relationship allowed them to conceive of variables as parameters without changing the properties of that function. For example, in a three variable function some students might conceive of treating one variable as a constant as creating an entirely new two-variable function. Instead, by keeping in mind the invariance of the relationship between variables, they could conceive of parameterizing any number of variables while still thinking about the same function. This way of thinking and the associated use of parameterization also allowed the students to generalize their understanding of function behavior to functions of any number of variables.

References

Harel, G. (2008a). DNR perspective on mathematics curriculum and instruction, Part I: focus on proving. *ZDM Mathematics Education, 40*, 487-500.

Harel, G. (2008b). A DNR perspective on mathematics curriculum and instruction. Part II: with reference to teacher's knowledge base. *ZDM Mathematics Education, 40*, 893-907.

Harel, G. (2008c). What is mathematics? A pedagogical answer to a philosophical question. *Current issues in the philosophy of mathematics from the perspective of mathematicians*, 265-290.

Martinez-Planell, R., & Trigueros, M. (2012). Students' understanding of the general notion of a function of two variables. *Educational Studies in Mathematics*.

Yerushalmy, M. (1997). Designing representations: Reasoning about functions of two variables. *Journal for Research in Mathematics Education, 28*(4), 431-466.

135

STUDENTS' AND EXPERTS' WAYS OF REASONING ABOUT PARTIAL DERIVATIVES ACROSS STEM CONTEXTS

Eric Weber
Oregon State University

Tevian Dray
Oregon State University

Corinne Manogue
Oregon State University

Mary Bridget Kustusch
DePaul University

David Roundy
Oregon State University

A common feature across STEM disciplines is the study of change. Mathematically, we express the concept of changing one parameter while fixing others by using partial derivatives. However, how we use partial derivatives and how we talk about partial derivatives vary dramatically across STEM disciplines. The purpose of this poster is to share our preliminary results from student and expert problem-solving interviews about partial derivatives.

Keywords: partial derivative, novice and expert, problem solving, STEM, clinical interview

A common feature across STEM disciplines is the study of change, whether studying how changing a design parameter affects the operation of a prototype, or how pressure changes when we adiabatically compress a gas. Indeed, the nature of scientific measurement is to control some physical quantities while measuring others. Mathematically, we express the concept of changing one parameter while fixing others by using partial derivatives. However, how we use partial derivatives and how we talk about partial derivatives vary dramatically across STEM disciplines. The purpose of this poster is to share our preliminary results from student and expert problem-solving interviews about partial derivatives.

We have found that many students—even those with a strong mathematics background—find partial derivatives particularly difficult. Further, in pilot interviews, we have found that experts employ a wide variety of reasoning strategies about partial derivatives, including the use of difference quotients constructed from numerical data, graphical reasoning about the slope at a single point, graphical reasoning about the shape of the graph, and symbolic computation. Even the experts we interviewed tended to get caught up in a single reasoning strategy, and did not necessarily transition spontaneously to other strategies.

The project we share in the poster involves two major strands of research activity, which focus on studying how STEM workers, from novice to expert, understand and use partial derivatives. The first strand of research activity is a survey across STEM disciplines of representations of partial derivatives used by experts, which informs our research and curriculum development efforts. We have used this research to identify normative practices that are common across many STEM disciplines, as well as those which are discipline specific. Based on these data, we are in the preliminary stages of developing learning trajectories intended to facilitate the evolution of students into professional users of partial derivatives in their own particular field. The second strand of research activity has involved a study of student understanding of partial derivatives across STEM courses, including physics, engineering and mathematics.

THE DUALITY PRINCIPLE AND LEARNING TRAJECTORIES IN MATHEMATICS EDUCATION

Eric Weber
Oregon State University
Eric.Weber@oregonstate.edu

Elise Lockwood
Oregon State University
elise314@gmail.com

The purpose of this paper is to argue that attention to students' ways of thinking should complement a focus on students' understanding of specific mathematical content, and that attention to these issues can be leveraged to model the development of mathematical knowledge over time using learning trajectories. To illustrate the importance of ways of thinking, we draw on Harel's (2008a, 2008b) description of mathematical knowledge as comprised of ways of thinking and ways of understanding. We use data to illustrate the explanatory and descriptive power that attention to the duality of ways of understanding and ways of thinking provides, and we propose suggestions for constructing learning trajectories in mathematics education research.

Keywords: Ways of thinking, Ways of understanding, Duality, Mathematical knowledge, Learning trajectories

Introduction

Learning trajectories (Simon & Tzur, 2004), which model how students' learning might develop over time, are becoming increasingly prevalent in mathematics education research. To date, much of the current research on learning trajectories focuses on the learning of particular mathematical topics such as fractions (Simon & Tzur, 2004), measurement (Gravemeijer, Bowers, & Stephan, 2003), multivariable functions (Weber, 2012), and trigonometry (Moore, 2012). While these are certainly valuable mechanisms that shed light on students' learning, we are not alone in observing that there is an aspect of learning (beyond student's knowledge of content) to which such learning trajectories do not currently attend. Empson (2011) stated:

> Most, if not all, current characterizations of learning trajectories do not address the practices that engender the development of concepts – although it's worth thinking about alternative ways to characterize curriculum standards and learning trajectories that draw teachers' attention to specific aspects of students' mathematical practices as well as the content that might be the aim of that practice (p. 573)

We agree with Empson's assessment and believe that there might be value in considering mathematical practice in learning trajectories, in addition to mathematical content. The purpose of this paper is to argue that attention to students' ways of thinking should complement a focus on students' understanding of specific mathematical content, and that attention to these issues can be leveraged to model the development of mathematical knowledge over time using learning trajectories (Simon & Tzur, 2004). We draw on Harel's (2008a, 2008b, 2008c) duality principle, which characterizes mathematical knowledge as comprised of *ways of thinking* and *ways of understanding* (defined momentarily) to propose a means by which learning trajectories can attend to both ways of thinking and ways of understanding.

Our theoretical contribution is to suggest that learning trajectories could incorporate and attend to mathematical practices as well as mathematical content. While Harel's duality principle and learning trajectories are both well-established in the mathematics education literature, we have found that combining these two existing theories provides meaningful language and perspectives to frame our ideas. We are not aware of other researchers who have made the same argument with any one existing framework, and we contend that our particular perspective provides a novel theoretical view.

Part 1: DNR Framework and Mathematical Knowledge

Harel (2008a, 2008b, 2008c) proposed the DNR based instruction framework as a way to think about the practice and teaching of mathematics. The constituent parts of the framework are the *duality* (D), *necessity* (N) and *repeated reasoning* (R) principles, which together comprise effective and meaningful mathematics instruction; in this paper we focus on the duality principle.

The *duality principle* states that mathematical knowledge consists of *both* students' understanding of particular content in mathematics and their thinking about the practice of doing mathematics. To clarify this distinction, Harel (2008a) introduced the notion of a *mental act*, which includes activities like interpreting, conjecturing, explaining, searching, and problem solving (p. 3). These mental acts are at the core of Harel's fundamental distinction between *ways of understanding* and *ways of thinking*. He proposed that a *way of understanding* is "a particular cognitive product of a mental act carried out by an individual" (p. 4) (Figure 1). He described a *way of thinking* as 'a cognitive characteristic of a person's ways of understanding associated with a particular mental act" (Harel, 2008a). Harel's analogy was that ways of understanding correspond to subject matter knowledge and ways of thinking correspond to conceptual tools.

As examples of the distinction in Harel's duality principle, we consider the mental acts of proof and problem solving. A particular proof of a given statement is a way of understanding that comes out of the mental act of proving, whereas a proof scheme is a way of thinking that characterizes the mental act of proving. In problem solving, a solution to a particular problem represents a way of understanding, but a general problem solving strategy, applicable across a variety of problems, suggests a way of thinking. Figure 1 shows a diagram of these terms' interaction.

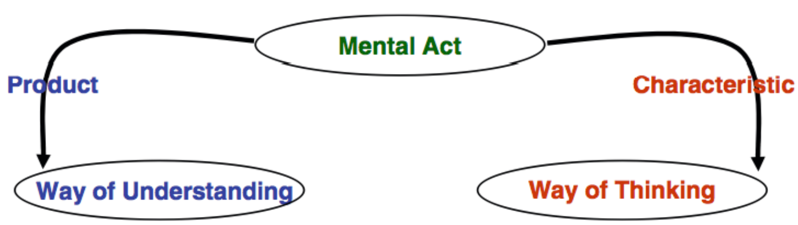


Figure 1: The mental act, way of understanding and way of thinking (Harel, 2008a).

Another key assumption of Harel's model of mathematical knowledge is that thinking and understanding affect each other. That is, "students develop ways of thinking only through the construction of ways of understanding, and the ways of understanding they produce are determined by the ways of thinking they possess" (Harel, 2008a). Thus, ways of thinking, applied in one moment to a novel situation, become a way of understanding. At the same time, the ways of thinking students develop occur from patterns observed in ways of understanding. This feedback between ways of thinking and ways of understanding is the core of Harel's duality principle. Based on this duality principle, we have found it useful to think about the development of mathematical knowledge at two levels: understanding and thinking. As a result, we argue that any representation of learning that focuses on the development of mathematical knowledge over time should take into account both aspects of that development. In order to frame our argument, we briefly discuss core elements of learning trajectories.

Part 2: A Brief Introduction to Learning Trajectories

When we describe learning trajectories, we mean representations (either predictive or descriptive) of the development of students' mathematical knowledge over time, most often in the context of specific tasks. Simon and Tzur (2004) first identified a hypothetical learning trajectory (HLT) as a model of how students' learning might occur over a period of time, with particular attention paid to students' mathematical activity and the role of tasks in engendering that activity. They proposed four principles for the hypothetical learning trajectory construct (Simon & Tzur, 2004, p. 93): 1) Generation of an HLT is based on understanding of the current knowledge of the students involved; 2) An HLT is a vehicle for planning learning of particular mathematical concepts; 3) Mathematical tasks provide tools for promoting learning of particular mathematical concepts and are, therefore, a key part of the instructional process; and, 4) Because of the hypothetical and inherently uncertain nature of this process, the teacher is regularly involved in modifying every aspect of the HLT. Consistent with Simon's model, most learning trajectories in mathematics education research elucidate how students might engage with and reflect on tasks, and, as a result, develop the mathematical understandings that we intend (Weber, 2012; Weber, Tallman, Byerley, & Thompson, 2012). Because it is impossible to account for all variation in an individual, HLTs are by their nature (and definition) hypothetical. These learning trajectories model the development of mathematical knowledge over periods of time ranging from a single lesson to entire grade levels.

We now can reformulate our argument, given the introduction of specific terminology. In this paper, we seek to leverage Harel's work (Harel, 2008a, 2008b, 2008c; Harel & Koichu, 2010), particularly his characterization of *duality* in terms of *ways of thinking* and *ways of understanding*, as we propose recommendations for constructing and evaluating learning trajectories. Ultimately we suggest modifications for the development of learning trajectories that attend to and incorporate duality.

Part 3: Importance of Attending to Duality: Examples from Data

In this section, we demonstrate how Harel's distinction between ways of thinking and ways of understanding can expand on typical descriptions of content knowledge. Though we present one example due to space limitations, our claims are based on multiple examples in a variety of content areas. The point of this example is provide some motivation and rationale for our belief that learning trajectories might attend to both *ways of thinking* and *ways of understanding*. In the following discussion, it is important to keep in mind the distinction between mental acts and ways of thinking. As Harel (2008c) points out (p. 3), mental acts include activities like proving, explaining, generalizing, and justifying. Ways of understanding are products of such mental acts, while ways of thinking are cognitive characteristics of them (p. 4). The example presented here focuses on the mental act of problem solving, and we highlight how duality (attention to both ways of thinking and ways of understanding) might offer some novel insight about a student's reasoning and activity.

The example presented below draws on Lockwood's (2013, in press) work with post-secondary students who solved advanced counting problems in videotaped, semi-structured interviews. This example shows how a student's way of thinking can span multiple ways of understanding particular problems, and how considering duality can provide novel insights about a student's learning. The point in presenting the example here is to emphasize a problem-solving approach (solving smaller, similar problems) as a way of thinking, and to show how that way of thinking affected both the student's ways of understanding particular problems and the researcher's interpretation of the students' combinatorial thinking. This data demonstrates a distinction Harel (2008a) himself made between instances of solving a problem and broader problem-solving approaches. Indeed, he identifies "looking for a simpler problem" (p. 6) as an example of a way of thinking about the mental act of problem solving.

The way of thinking of solving smaller, similar problems is demonstrated across two problems for a particular student, Anderson. While we cannot provide mathematical detail due to space, we briefly describe his use of smaller cases on each problem, arguing that his use of the same strategy across multiple problems suggests a way of thinking for him. Additionally, in his work on these problems, Anderson displayed a way of thinking that would feed into his understanding of a particular problem, and we discuss this interaction below.

The Passwords problem states, *A password consists of eight upper case-letters. How many such 8-letter passwords contain at least three Es?* In this episode, Anderson was trying to decide which of two different expressions was correct, and he went to a smaller case (a 4-letter word containing three Es) to decide which made sense. He found a numerical discrepancy even in the smaller case, and he was able to reason about the smaller case to determine the correct answer to the problem. Anderson's use of the smaller, similar problem was a vital part of him successfully evaluating the alternative solution and determining an accurate answer. After this episode, the interviewer asked Anderson to reflect on his use of the smaller case. In his reflection, it seems that he had anticipated how working with the smaller problem would facilitate his manipulation of the passwords, suggesting to us that the strategy might represent a way of thinking for him.

Later in the interview, Anderson solved the Groups of Students problem: *In how many ways can you split a class of 20 into four groups of five?* In this problem, Anderson first tried to reduce the problem to splitting eight students into four groups of two, but this quickly became unwieldy for him to handle. Recognizing this difficulty, he then reduced the number of groups and the total number of students to make the problem more tractable. He divided a class of four into two groups, and through systematic listing found that there were three ways to do this. He then attempted to determine how a class of six could be split into two groups and found that there were 10 such possibilities. Anderson continued in this way, he made an initial guess at what the general formula might be: "the number of students choose the size of the groups, divided by the number of groups." We note that this formula is incorrect, but given his work it is a reasonable first attempt. Recognizing that he wanted to test out this guess at a formula, Anderson proceeded to solve another smaller problem, this time splitting six students into three groups of two. He wrote out solutions and similarly developed a pattern, continuing to reason about the problem. We ultimately ran out of time for Anderson to come up with a correct solution on his own, but his work with the smaller problems proved fruitful for him, and he was able to make sense of the correct answer when it was presented to him.

In our analysis of Anderson's work, we infer that Anderson's ways of understanding in the two problems fed back into a particular way of thinking (solving smaller, similar problems). We contend that the use of multiple smaller problems and the emergence of patterns supplemented and expanded his previous way of thinking about the Passwords problem. Specifically, we argue that as a result of his work on the Groups of Students problem, Anderson might have learned that he had to be strategic in his choice about how to reduce the problem – simply reducing any parameters might not be helpful. This is seen in his first unsuccessful attempt at breaking a group of eight into four groups of two, and this is an insight that might not have arisen had he only solved the Passwords problem. We would thus argue that Anderson's way of thinking (the strategy of solving smaller problems) is more robust because of the ways of understanding with which he engaged on the two problems.

From the perspective of examining the content that Anderson learned, Lockwood (2013, in press) has made the case for combinatorial implications of this work. However, we emphasize that beyond the content, these two episodes together reveal an important aspect of Anderson's learning – his use of a particular problem-solving approach – that can be described in terms of a way of thinking. Even more, Anderson's work on both of these problems provides evidence of the duality principle at play in his problem solving, suggesting that his ways of thinking and ways of understandings reflexively interacted as he solved counting problems.

While this is just one brief example, it is meant to show that if researchers seek to articulate aspects of students' learning via learning trajectories, there could be value in targeting both ways of understanding ways of thinking such as those that Anderson displayed. By emphasizing this duality, we suggest that the interaction between ways of understanding and ways of thinking actually shed light on Anderson's combinatorial conceptions, and these provide explanatory aspects of his work that would otherwise not come up. By observing his way of thinking, ways of understanding, and how they interact across multiple problems, we have a more complete picture of how he may think about and learn counting problems.

Part 4: Incorporation of Duality in Designing Learning Trajectories

The purpose of this section is to provide specific recommendations for ways in which attention to the duality principle could shape how researchers think about and use learning trajectories. We suggest that researchers might recognize the potential that explicit attention to the duality of ways of thinking and ways of understanding might shape the understandings we might expect students to develop. To frame our recommendations, and to identify in what ways we see the focus on duality contributing to the current notion of a learning trajectory, we again consider Simon & Tzur's (2004) elements of a hypothetical learning trajectory. Under each element, we consider what a focus on duality contributes, and how researchers might practically focus on duality in the construction and revision of learning trajectories.

Table 1. Incorporating Duality into Learning Trajectories

Principle	**Considerations Duality Introduces**	**Recommendations for Researchers**
Generation of an HLT is based on understanding of the current knowledge of the students involved	Understanding the current knowledge of students involved might entail a model of the students' ways of thinking and their ways of understanding that is grounded in the literature base and develops from interactions with the students.	Consider that difficulties or insights students appear to have could be related to their ways of thinking as well as their understanding of particular content. Ask questions across a variety of problems and domains to understand if their difficulties are content specific or involve ways of thinking.

An HLT is a vehicle for planning learning of particular mathematical concepts	An HLT is a vehicle for anticipating the development of mathematical knowledge, comprised of both ways of thinking and ways of understanding. The development of this knowledge can be represented as a feedback loop between ways of thinking and ways of understanding.	Explicitly articulate mental acts (such a problem solving, justifying, proving, explaining, etc.) and anticipate potential products of (ways of understanding) and characterizations of (ways of thinking) mental acts that might arise for students.
Mathematical tasks provide tools for promoting learning of particular mathematical concepts and are, therefore, a key part of the instructional process	Mathematical tasks provide a means to engender learning of mathematical concepts, and in doing so, provide a means to affect the development of ways of thinking. In turn, mathematical tasks that focus on engendering ways of thinking spur the development of particular subject matter knowledge.	Create tasks that help gain insight into students' ways of thinking by creating opportunities in tasks for students to reflect on their approaches and solutions *across* problems and situations. Be aware that certain mathematical domains may more or less effectively facilitate particular ways of thinking (e.g. solving smaller simpler problems)
Because of the hypothetical and inherently uncertain nature of this process, the teacher-researcher is regularly involved in modifying every aspect of the HLT	The teacher-researcher should consider modifications to the learning trajectory with both ways of thinking and ways of understanding in mind, developing and refining activities intended to engender those ways of thinking and ways of understanding.	Document the development of students' content knowledge in conjunction with their ways of thinking, explicitly attending to how these two aspects of students' mathematical knowledge interact.

We want to emphasize three things about the recommendations above. First, our recommendations should not be considered as a call to do a new "type" of learning trajectories. Instead, we have provided suggestions for what incorporating duality into learning trajectories might entail. Second, our recommendations should not be considered an exhaustive list. We include these points as a means to drive discussion about some major issues to consider with the inclusion of ways of thinking in learning trajectories. Third, content plays a significant role in how ways of thinking develop, and certain domains may be more appropriate than others for fostering specific ways of thinking. For instance, combinatorics is a particularly powerful context for thinking of solving smaller, similar problems, as the *need* for such work with smaller cases can easily be motivated in a setting that typically deals with very large and unwieldy numbers. While this way of thinking may be effectively developed in a domain like combinatorics, it could be further refined and developed in other mathematical areas, each of which might elicit different aspects of the way of thinking.

Conclusion

In this paper, we have suggested that there might be an additional aspect of mathematical learning that could be incorporated into learning trajectories. In particular, Harel's duality principle gives us language for articulating what we have noticed in our own work – that mathematical knowledge might include both mathematical content and mathematical practices.

References

Empson, S. B. (2011). On the idea of learning trajectories: Promises and pitfalls. *The Mathematics Enthusiast, 8*(3), 571-596.

Gravemeijer, K., Bowers, J., & Stephan, M. (2003). Chapter 4: A Hypothetical Learning Trajectory on Measurement and Flexible Arithmetic. *Journal for Research in Mathematics Education. Monograph, 12*, 51-66.

Harel, G. (2008a). DNR perspective on mathematics curriculum and instruction, Part I: focus on proving. *ZDM Mathematics Education, 40*, 487-500.

Harel, G. (2008b). A DNR perspective on mathematics curriculum and instruction. Part II: with reference to teacher's knowledge base. *ZDM Mathematics Education, 40*, 893-907.

Harel, G. (2008c). What is mathematics? A pedagogical answer to a philosophical question. *Current issues in the philosophy of mathematics from the perspective of mathematicians*, 265-290.

Harel, G., & Koichu, B. (2010). An operational definition of learning. *The Journal of Mathematical Behavior, 29*(3), 115-124.

Lockwood, E. (2013). *Developing facility with sets of outcomes by solving smaller, simpler counting problems.* Paper presented at the 16th Annual Conference on Research in Undergraduate Mathematics Education, Denver, CO.

Lockwood, E. (in press). The strategy of solving smaller, simpler problems in the context of combinatorial enumeration. *Research in Collegiate Mathematics Education*.

Simon, M., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning, 6*(2), 91-104.

Weber, E. (2012). *Students' Ways of Thinking about Two Variable Functions and Rate of Change in Space*. doctoral dissertation. Arizona State University. Tempe, AZ.

Weber, E., Tallman, M. A., Byerley, C., & Thompson, P. W. (2012). Introducing derivative through the calculus triangle. *Mathematics Teacher, 106*(4), 274-278.

137

WHAT IS A PROOF? A LINGUISTIC ANSWER TO AN EDUCATIONAL QUESTION

Keith Weber
Rutgers University

Proof is a central concept in mathematics education, yet mathematics educators have failed to reach a consensus on how proof should be conceptualized. I advocate defining proof as a clustered concept, in the sense of Lakoff (1987). I contend that this offers a better account of mathematicians' practice with respect to proof than previous accounts that attempted to define a proof as an argument possessing an essential property, such as being convincing or deductive. I also argue that it leads to useful pedagogical consequences.

Key words: Cluster model; linguistics; proof

It is widely accepted that having students successfully engage in the activity of proving is a central goal of mathematics education (NCTM, 2000; Harel & Sowder, 1998). Yet mathematics educators cannot agree on a shared definition of proof (Balecheff, 2002; Reid & Knipping, 2010). This is recognized as problematic: without a shared definition, it is difficult for mathematics educators to meaningfully build upon each another's research and it is impossible to judge if pedagogical goals related to proof are achieved (e.g., Balacheff, 2002; Weber, 2009). Until now, most mathematics educators have sought to define proof as an argument possessing one or more desirable properties (e.g., being deductive or being convincing) without a consensus on which property or properties are the essence of proof. The main thesis of this paper is that viewing proof as a clustered model in the sense of Lakoff (1987) offers a better account of how proof is practiced by mathematicians.

Previous attempts to characterize proof

Through most of the 20th century, the philosophy of mathematics was dominated by the study of logic and the foundations of mathematics. Proof was defined as a formal object: a proof was situated in a formal language with explicit rules for well-formed formulae and logical inference; a sequence of well-formed formulae was a proof if each formula was a premise or the result of applying a rule of inference to previous formulae. While this characterization has the virtue of being an objective method for identifying a proof, it had the drawback of bearing little relevance to mathematical practice. It is rare for a mathematician to produce a proof that satisfies these criteria.

More recently, philosophers and mathematics educators have sought to characterize the proofs that mathematicians actually produce. In this line of research, it is typical for a researcher to propose a property (or occasionally multiple properties) that arguments must satisfy to be a proof. Inevitably, another researcher will challenge this definition by producing a counterexample that either satisfies these properties but is not a proof or is a proof but fails to satisfy these properties. For instance, Hanna (1995) argued that computer-assisted proofs (e.g., Appel and Haken's proof of the four-color theorem) challenge the notion that a proof is open to public inspection. Similarly, Larvor (2012) dismissed the notion that proofs being non-ampliative (i.e., not supplying conclusions that exceed the premises or being deductive) as he claimed that some authoritative arguments are non-ampliative but not proofs. Harel and Sowder (1998) defined a mathematical proof as an argument that convinces a mathematician, but others have argued there are convincing arguments that do not constitute proofs (e.g., Tall, 1989). (For instance, Eccheveria (1996) said the mathematical community is convinced that the unproven Goldbach's Conjecture is true based on empirical evidence).

I believe that the reason that mathematics educators have collectively failed at the problem of defining proof is due to a faulty assumption of what a solution to this problem must look like. The search is for a set of properties that will distinguish whether an argument is a proof. It is natural to attempt to define proof this way. This is, after all, how mathematical categories are defined (Alcock & Simpson, 2002). However, I claim that a search for this type of definition is doomed for failure.

Aberdein (2009) coined the term, "proof*", as "species of alleged 'proof' where there is no consensus that the method provides proof, or there is a broad consensus that it doesn't, but a vocal minority or an historical precedent point the other way". As examples of proof*, Aberdein included "picture proofs*, probabilistic proofs*, computer-assisted proofs*, [and] textbook proofs* which are didactically useful but would not satisfy an expert practitioner". This presents a paradox for creating a definition that lists properties that distinguish proofs from non-proofs. Take picture proofs*, for instance. Such a definition would either permit some convincing picture proofs* to count as proofs or it would say no picture proofs* were proofs. If the former occurred, one could rebut the definition by citing the large number of mathematicians who said these picture proofs* were not real proofs. If the latter occurred, one could rebut the definition by citing the large number of mathematicians (or at least the vocal minority) who disagree. Similar arguments could be made for all types of proofs*. Of course, the proponent of the definition could resort to saying that the mathematicians who disagree with his or her perspective are mistaken (e.g., computer-assisted proofs* are proofs but probabilistic proofs* and picture proofs* are not and any mathematician who disagrees is in error). However, aside from making the dubious argument that a considerable number of mathematicians do not really understand their craft, this would challenge the claim that his or her definition was descriptive of mathematical practice.

Lakoff's clustered concepts

Lakoff (1987) noted that "according to classical theory, categories are uniform in the following respect: they are defined by a collection of properties that the category members share" (p. 17). As noted above, this assumption underlies the ways in which most philosophers and mathematics educators seek to define proof. However, Lakoff's thesis is that most real-world categories cannot be characterized this way. In particular, he argued that some categories might be better thought of as *cluster models*, which he defined as occurring when "a number of cognitive models combine to form a complex cluster that is psychologically more basic than the models taken individually" (p. 74).

As an illustrative example, Lakoff considered the category of *mother*. To Lakoff, there are several types of mothers, including the birth mother, the genetic mother, the nurturance mother (i.e., the adult female caretaker of the child), and the marital mother (i.e., the wife of the father). In the prototypical case, these concepts will converge-- that is, the birth mother will also be the genetic mother, the nurturance mother, and so on. And indeed, when one hears that the woman is the mother of a child, the default assumption is that she assumes all of these roles.

Lakoff raised two other points that will be relevant to this paper. First, there is a natural desire to pick out the "real" definition of mother, or the true "essence" of mother. However, this does not seem possible. Different dictionaries list different conceptions of mother as their primary definition. Further, sentences such as, "I was adopted so I don't know who my *real mother* is" and "I am uncaring so I doubt I could be a *real mother* to my child" both are intrinsically meaningful yet define *real mother* in different ways. Second, in cases where there is divergence in the clustered concept of mother (e.g., a genetic but not adoptive mother), compound words exist to qualify the use of mother. Calling one a birth mother typically indicates that she in not the nurturance mother; calling one an adoptive mother or a stepmother indicates that she is not the birth mother.

Proof as a clustered concept

The main thesis in this paper is that proof can be productively characterized as a cluster model. Exactly what cognitive models compose this cluster can be the subject of interesting debate. What I present is a first attempt at categorizing proofs:

- *A proof is an argument that is deductive and non-ampliative argument:* Each statement in a proof should be a premise or a necessary logical consequence of previous statements.
- *A proof is an argument that would convince a contemporary mathematician who knew the subject* (adapted from Davis & Hersh, 1981)
- *A proof is an argument in a natural language and symbolic representation system where there are socially sanctioned rules of inference*. It's noteworthy that some researchers view a proof as an argument that can be translated into a formal proof (e.g., Azzouni, 2004; Mac Lane, 1986). If so, the distance between the language of formal proofs and the proofs that are actually produced should not be too great.
- *A proof is an argument that convinces a particular community at a particular time* (adapted from Balacheff, 1987). This definition emphasizes the social role of proof and situates proof as dependent on time and culture.
- *A proof is an argument that is a blueprint that knowledgeable mathematicians can use, in principle, to write a complete proof with no logical gaps.*

My claim is that an argument that satisfies all these properties would be judged by (nearly) all knowledgeable mathematicians as constituting a proof and an argument that satisfied none of these conditions would be rejected by (nearly) all mathematicians as a non-proof. Arguments that satisfied some, but not all of these conditions, such as Aberdein's (2009) proofs*, would be regarded as controversial. Further, it would be desirable, if possible, to improve these arguments so they satisfied the properties that they lacked.

The usual stratagem of rejecting a conception of proof-- citing a counterexample where the conception did not reflect reality-- would not apply here. In this characterization, there is no guarantee that every proof satisfies every property. However, this conception of proof does make some concrete predictions about proof that can be tested. First, arguments satisfying all the criteria above would be considered more "proof-like" or representative of proof than those that would not. A proof in knot theory that relied on diagrams or kinesthetic motion (see Larvor, 2012 or Rav, 1999) might be accepted as a proof, but the mathematical community at large would consider these proofs unusual in this respect. Second, if a mathematician was told a proof of a particular theorem existed, his or her default assumption would be that the argument satisfied each of the constraints, even though he or she would be aware that this might not necessarily be so. Third, there should be compound words that qualify proof-like arguments that satisfy some, but not all, of these criteria. Indeed, Aberdein (2009) provided a list including picture proofs* (convincing deductive arguments in a non-symbolic/natural language representation system) and probabilistic proofs* (arguments that are arguably convincing but not deductive). Fourth, if mathematicians were asked what the true essence of a proof was, they should not all list one of the criteria above. Rather, their responses should be heterogeneous (i.e., the essence of proof varies by mathematician) or many should say the question is unanswerable. That Aberdein (2009) alleged that his proofs* have both their adherents and detractors is evidence toward this point.

Benefits of this conception of proof

Previous attempts to define proof

Previous attempts to define proof in mathematics education have usually latched on to one or two of the properties above and treated it as the sole criterion on which to judge proofs. Harel and Sowder (1998) defined mathematical proof in terms of arguments that convince mathematicians, Hoyles and Kucheman (2002) treated proofs as arguments that are deductive, Weber and Alcock (2009) defined proof as arguments within a representation system, and Balacheff (1987) defined a proof as time- and community-dependent. Stylianides (2007) desired arguments to be both deductive and in an age-appropriate representation system to constitute a proof. Each definition essentially says the true essence of proof lies in one or two of the attributes described above, treating the other listed criteria as either tangential or a consequence of a proof's essence. For instance, Hanna (1991) referred to the formality of proof, or placing the argument in a representation system, as merely a "hygiene factor" and Harel and Sowder (1998) believed the fact that mathematical proofs are deductive is a necessary consequence of these proofs convincing mathematicians. As noted earlier, such characterizations cannot account for mathematicians' practice with respect to proofs* and thus seem not to categorize mathematical practice.

More appropriate pedagogical suggestions

At a broad level, the components of the clustered model of proof are correlated with one another. For instance, in general, as an argument becomes more deductive, it tends to become more convincing, easier to translate into a formal proof, and more likely to be sanctioned by one's peers. Hence, encouraging students to make their arguments more deductive would usually make their arguments more proof-like in other respects as well. However, this is not the case if we take some of these criteria to extremes.

For a first example, suppose we strive to present students with arguments that are as convincing as possible in geometry. In many cases, an exploration on a dynamic geometry package would be extremely convincing, both for mathematicians and for students (de Villiers, 2004). For a student, such explorations would probably be *more* convincing than a complicated deductive argument because the student may worry that he or she has overlooked an error in the argument. If we view the mode of reasoning (deductive vs. perceptual) and the representation system in which an argument is couched as irrelevant, it is difficult to argue why demonstrations on dynamic geometry software packages are not proofs.

A similar claim relates to how formal an argument is. Increasing the formality of an argument usually makes the argument more deductive and more acceptable to the mathematical community. However, it is generally accepted that there is a point where an argument is "formal enough" and making it more rigorous would be detrimental. An argument with large logical gaps might be judged as not convincing and not accepted by mathematicians. However, filling in *all* the gaps would make the proof impossibly long and unwieldy. The result would be a proof that masks its main ideas. As understanding these ideas is important for determining the validity of the proof, increasing the rigor of the proof would lessen its persuasive power.

If we want students and teachers to present proofs that satisfy all or most of the criteria above, it would be best not to focus on a single criterion. Not only would the other criteria be ignored, a singular focus on one criterion might actually lessen the possibilities of the other criteria being achieved.

References

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science*, 14, 1-8.
Alcock, L. & Simpson, A. (2002). Definitions: Dealing with categories mathematically. *For the Learning of Mathematics*, 22(2), 28-34.
Azzouni, J. (2004). The derivation-indicator view of mathematical practice. *Philosophia Mathematica*, *12*, 81-106.
Balacheff, N. (1987). Processus de preuves et situations de validation. *Educational Studies in Mathematics*, 18(2), 147-176.
Balacheff, N. (2002). The researcher epistemology: A deadlock from educational research on proof. In F.L. Lin (ed.) *2002 International Conference on Mathematics Education-Understanding proving and proving to understand*. NCS and NSUT: Taipei.
Davis, P. J. and Hersh, R. (1981). *The mathematical experience*. New York: Viking Penguin Inc.
de Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras, 24*, 17-24.
de Villiers, M. (2004). The role and function of quasi-empirical methods in mathematics. *Canadian Journal of Science, Mathematics, and Technology Education*, 4, 397-418.
Echeverria, J. (1996). Empirical methods in mathematics. A case study: Goldbach's Conjecture. In G. Munevar (ed) *Spanish studies in the philosophy of science*,. (pp. 19-56). Dordrecht: Kluwer
Hanna, G. (1991). Mathematical proof. In D. Tall (ed.) *Advanced Mathematical Thinking.* Dordrecht: Kluwer.
Hanna, G. (1995). Challenges to the Importance of Proof. *For the Learning of Mathematics,* 15(3), 42-50.
Harel, G. & Sowder, L. (1998). Students proof schemes. *Research in Collegiate Mathematics Education, 3*, 234-282.
Hoyles, C. & Kuchemann, D. (2002). Students' understanding of mathematical implication. *Educational Studies in Mathematics*, 51, 193-223.
Lakoff, G. (1987). *Women, fire, and dangerous things.* Cambridge University Press: Cambridge.
Larvor, B. (2012). How to think about informal proofs. *Synthese*, 187, 715-730.
Mac Lane, S. (1986). *Mathematics: Form and function*. Springer: Berlin.
National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.
Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica*, *7*, 5-41.
Reid, D. & Knipping, C. (2010). *Proof in mathematics education*. Sense Publishers: Rotterdam.
Stylianides, A. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 38, 289-321.
Tall, D. (1989). The nature of mathematical proof. Mathematics Teaching, 127, 28-32.
Weber, K. (2009). Review of Boero's Theorems in schools. Mathematical Thinking and Learning, 11, 289-294.
Weber, K. & Alcock, L. J. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D.A. Stylianou, M.L. Blanton & E. Knuth (Eds.), Teaching and Learning Proof Across the Grades: A K-16 Perspective (pp. 323-338). New York, NY: Routledge.

EXPLORING STUDENTS' WAYS OF THINKING ABOUT SAMPLING DISTRIBUTIONS

Aaron Weinberg
Ithaca College

The concept of a sampling distribution plays a central role in the process of making statistical inferences. However, students typically struggle to understand and reason about sampling distributions. This study seeks to characterize the ways undergraduate students think about sampling distributions in scenarios involving repeated sampling and making statistical inferences. Eight students in an introductory statistics class worked on problems involving sampling distributions during a semi-structured interview. A framework was developed based on their responses to describe the ways they discussed and coordinated various aspects of the population and sampling distributions by focusing on the processes of sampling and repeated sampling; these descriptions suggest that explicitly coordinating particular aspects of these processes may correspond to the robustness of students' conceptions of sampling distributions.

Key words: Statistics, Sampling Distributions, Sampling Variability, Conceptual Operations

Introduction

Statistical inference is the process by which conclusions about a population are drawn based upon evidence obtained from a sample of the population. Developing students' ability to make and understand statistical inferences is a key focus for most college-level introductory statistics courses. Yet, research suggests there are substantial gaps in students' informal and formal understanding of statistical inference (e.g., Zieffler et al., 2008).

In particular, students have difficulty understanding the concept of a sampling distribution, which describes the relative frequency of statistics we would expect to see when collecting all possible samples of a given size from a population. Sampling distributions connect probability models with statistical inference, enabling us to compare results from an observed sample with a theoretical distribution. Research has documented students' difficulties with the concept of sampling distributions (e.g., Chance, delMas & Garfield, 2004), and many educators may underestimate the complexity involved in building a coherent understanding of the concept.

The goal of this research project is to construct a framework that facilitates a fine-grained description of students' thinking about sampling distributions and to use the framework to describe some of the ways students in introductory statistics courses think about sampling distributions when solving problems involving inference and repeated sampling.

Background

Much of the research on students' understanding of sampling distributions has taken the form of instructional interventions using computer simulation methods (CSMs). Although many studies have argued that CSMs aid student understanding of sampling distributions, Mills (2002) noted that most of these studies did not report empirical data. In addition, most of the empirical studies have documented students' misconceptions and difficulties rather than providing descriptions of the ways students might come to understand the underlying concepts.

Saldanha and Thompson (e.g.; Saldanha & Thompson, 2002a; 2002b; 2006; 2007) have investigated students' conceptions of sampling distributions and understanding of sampling variability, focusing on students' imagery of a process of repeatedly sampling from a population. In several of their earlier studies, they found that the more successful students "had developed a multi-tiered scheme of conceptual operations centered around the images of repeatedly sampling from a population, recording a statistic, and tracking the accumulation of statistics as they distribute themselves along a range of possibilities" (Saldanha & Thompson, 2002b p. 261). Based on these studies, Saldanha and Thompson (2002b, p. 261) identified three levels of the sampling process:

- Level 1: Randomly select items to accumulate a sample of a given size from a population. Record a sample statistic of interest.
- Level 2: Repeat Level 1 process a large number of times and accumulate a collection of statistics.
- Level 3: Partition the collection in Level 2 to determine what proportion of statistics lie beyond (below) a given threshold value.

They found that successful students were able to clearly distinguish between these levels. In contrast, less-successful students often confounded the number of samples with the sample size and struggled to coordinate ideas across the various levels of the sampling process.

In subsequent studies, Saldanha and Thompson (2006; 2007) described the way students in a teaching experiment developed an understanding of the concept of a sampling distribution. The way the students engaged in the instructional intervention appeared to follow a three-phase trajectory:

- Phase 1: Students focus their attention on an individual sample as they select and aggregate items from the population. They use this sample to estimate a value for the parameter.
- Phase 2: Through the process of repeated sampling, students focus their attention on a collection of statistics and use the resulting distribution to estimate the parameter.
- Phase 3: Students repeat the repeated-sampling process to focus on comparing collections of values. They describe the similarities of the distributions and use this to discuss the "unusualness" of a particular sample statistic.

In addition to Saldanha and Thompson's framework, several researchers have suggested that students need to understand and coordinate numerous ideas about samples and the sampling (and resampling) process in order to understand sampling distributions (e.g., Bakkar & Gravemier, 2004; Chance, et al., 2004; Pfannkuch & Reading, 2006, Saldanha & Thompson, 2002b). Students must be able to understand and make comparisons between different samples and sampling distributions. They need to attend to the role played by sample size and understand the relationship between samples and the populations from which they are drawn. Students must be able to reason about distributions of data and to make use of proportional reasoning; reasoning about distributions requires the coordination of two or more attributes of a distribution, where the attributes are measures of center, spread, and shape. Students need to recognize what sampling variability represents and the role variability plays in the outcomes of a distribution. In addition, students need to be able to reason proportionally about the outcomes represented in the distribution in order to connect the sampling distribution to ideas of inference.

Taken together, these previous results suggest that a framework for describing the details of students' conceptions of sampling distributions should take into account the statistical processes (i.e. sampling and resampling) and objects (e.g., centers of distributions) that students attend to and coordinate as they engage in situations involving repeated sampling.

Methods

Eight undergraduate students participated in the study; they were all enrolled in a one-semester introductory statistics course with a target audience of mathematics majors and minors. The course was based on the CATALST curriculum (Garfield, delMas & Zieffler, 2012), which engages students in simulation and randomization methods to construct empirical sampling distributions from null and bootstrap models. Each student participated in two semi-structured interviews, one near the mid-point of the semester and the other near the end of the semester. In the interview, students were asked to describe what a sampling distribution is and then to work on four problems, which are outlined in Table 1.

Problem 1: You have a population of men with a mean height of 69 inches. In Study 1 you take a 5-person sample each day at a post office and record the mean height; in Study 2 the post office takes 50-person samples. In which study will you record more days over 71 inches?
Problem 2: Researchers collect many samples of 50 tires and compute the mean tread life of each; their results are displayed in a histogram [included in the question]. If you collect a new sample of 50 tires and find that it has a mean tread life of 6 years; is that evidence that the new tires last longer than the old ones?
Problem 3: Researchers sent out 2600 identical resumes; half had “white sounding” names and the others had “black sounding” names. They received 121 positive responses for the white-sounding names and 87 for the black-sounding names. Then they ran 500 simulations under the assumption that 208 out of every 2600 names should receive a response and the results are displayed in a histogram [included in the question]. Should the researchers be concerned?
Problem 4: Given a histogram of a population of test scores and histograms of four potential sampling distributions, which histogram(s) could represent sampling distributions with sample sizes of 4 and 50?

Table 1. Outline of interview questions

The students were asked to think aloud as much as possible while working on the problems. After working on each problem, the interviewer asked questions designed to challenge the student’s reasoning and conclusions. The interviews were video-recorded and transcribed.

Analysis of the data was conducted using grounded theory (Strauss & Corbin, 1990). The students’ utterances were read and categorized according to the mathematical and statistical processes and objects of their attention, the types of conclusions they drew, and their reasoning for drawing these conclusions. Then the utterances were re-coded axially, resulting in the categories described in the results.

Results and Analysis

All of the students were able to provide correct solutions to at least two of the four problems and were able to justify their reasoning using correct descriptions of the sampling distributions involved. All of the students were able to discuss the process of repeated sampling and the idea of partitioning a collection of sample statistics, suggesting that their reasoning was at Level 3 of Saldanha and Thompson’s (2002b) framework. In addition, all of the students were able to provide a correct solution to—and explain their reasoning on—problems 1 and 4 (which involved comparing sampling distributions) on at least one of their interviews; this suggests that they had reached Phase 3 of Saldanha and Thompson's (2007) hypothetical learning trajectory.

Although they appeared to be at similar levels according to these frameworks, the students still attended to different aspects of sampling and resampling as they described and interpreted sampling distributions. The different collections of concepts they coordinated, in turn, led to more or less robust and sophisticated ways of justifying their arguments. For example, consider the two students' arguments for why the small post office (in Problem 1) would record more days with a mean over 71 inches:

> Student 1: I think that the one with the larger sample size will see more consistent data in that their means will probably be relatively closer to each other every single day, whereas I think the impacts of the smaller sample size will shift that mean to extremes more often. So you'll see a larger—I guess a larger distribution in the statistics that you receive.

> Student 2: So, the small post office is a pretty small sample, so you can have, like, a pretty wide variety with just five people, but if you have 50 people, a lot of them... because 69 inches is the average, so you're saying a lot of people are 69 or close to 69, otherwise you wouldn't get that as the average. So if you have a group of 50, there's going to be a fair amount of them with around 69 inches, but if you only have 5 then you could have a really tall guy and you might have one or two near the average but you could also have some short guys, and it's just more likely that your average will be further from the 69 inches.

In his justification, Student 1was able to explicitly describe a process of repeated sampling and the relationship between the statistic (i.e. the sample mean) and the spread of the sampling distribution; he also alluded to a process of sampling by describing the role played by the size of the sample. Student 1's reasoning is summarized in Figure 1.

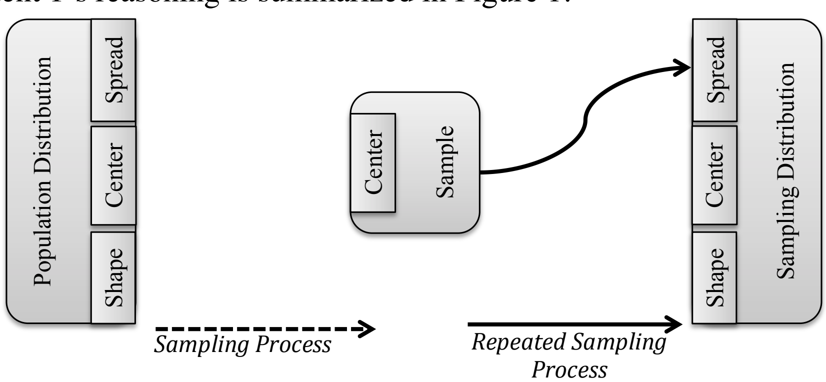


Figure 1. Student 1's reasoning

Student 2 was able to explicitly describe a process of sampling particular individuals (in a group of a specific size) from a population by attending to the center, spread, and—implicitly—the shape of the population (i.e. unimodal and symmetric); he also implicitly alluded to a process of repeated sampling (i.e. "more likely"), comparing the distance of the sample averages from the center of the population. Student 2's reasoning is summarized in Figure 2.

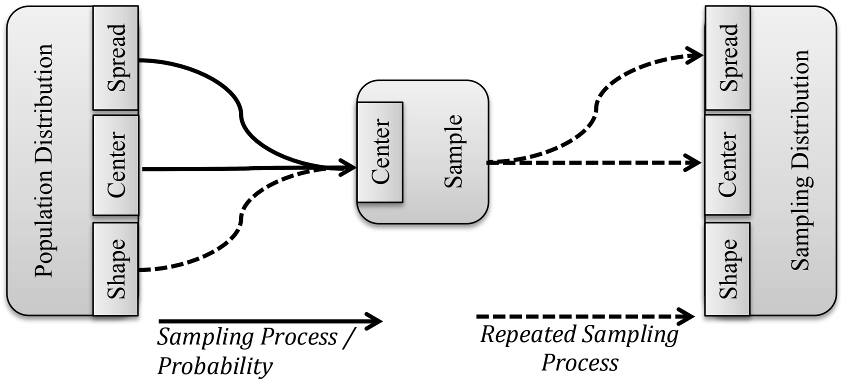


Figure 2. Student 2's reasoning

Student 1 had difficulty addressing the interviewer's subsequent questions (which were designed to probe his reasoning). For example, the interviewer asked Student 1 to discuss the effects of the composition of individual samples, noting that both large and small samples should have roughly the same proportion of "tall" people; the interviewer asked Student 1 to explain why this reasoning wouldn't imply that the two studies would have the same number of days with averages over 71 inches. Student 1 was unable to address this question effectively, replying:

> Student 1: Over 71 inches is now considered tall, okay. Umm, I mean... that makes sense. It would make sense if you would get one tall person, then you would see one in five people as 20% of your statistic as tall, that doesn't extrapolate to the likelihood of getting a tall person to walk through your door. That's just saying that 20% of the people you surveyed that day were tall. I think the odds of getting a tall person is the same.

In contrast, Student 2 was able to address the interviewer's subsequent questions by focusing on the sampling process and the way that was coordinated with the (implied) repeated sampling. For example, the interviewer suggested that a large sample would be more likely to include tall individuals and asked Student 2 to discuss why this wouldn't imply that the study with large samples would have more days with averages over 71 inches. Student 2 replied:

> Student 2: It is more likely that you will get a really tall or really short person, but you also have 49 other people that—a lot of them will be close to the average and then... you'll probably have a really tall guy and a really short guy that kind of average each other out.

These examples suggest that we can describe students' thinking about sampling distributions by identifying the aspects of (1) the population distribution, and (2) the sampling distribution that they explicitly and implicitly focus on, whether they leverage (3) the concept of the sample (and sample statistic) in their reasoning, and the way they coordinate these aspects by explicitly or implicitly describing (4) the sampling process and (5) the repeated sampling process. These myriad options can be visualized in Figure 3.

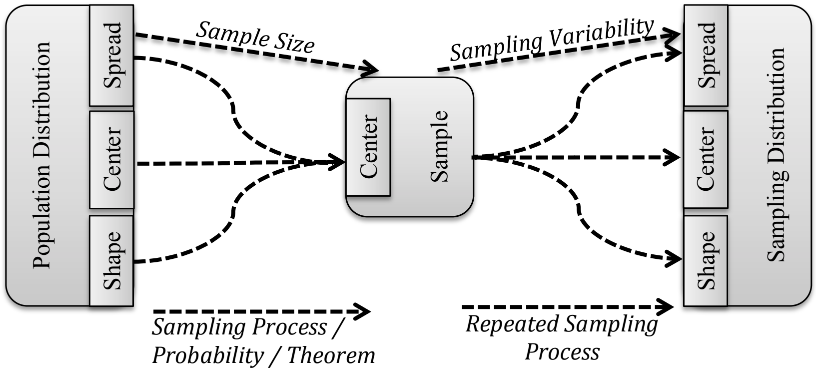


Figure 3. Sampling distribution framework for coordination of aspects

Each box in the diagram represents an aspect of a statistical object (e.g., the center of a population distribution) that can be attended to; the curved arrows represent the potential to coordinate the aspects. The "sampling process" and "repeated sampling process" arrows represent two processes students can leverage as mechanisms for coordinating the various aspects; the "sample size" and "sampling variability" are aspects of the two processes that can be attended to that influence the way the various aspects can be coordinated. In an individual student's reasoning, these aspects and processes may be attended to explicitly (represented by solid arrows) or implicitly (represented by dashed arrows).

Discussion

The framework presented here offers a way to document and categorize the various aspects of population and sampling distributions that students attend to as they work in scenarios involving repeated sampling and the resulting sampling distributions; in addition, it offers a way to identify which of these aspects students coordinate, the statistical processes they focus on as they coordinate the aspects, and the ways the ideas of sample size and sampling variability influence their thinking.

As shown in the examples from Student 1 and Student 2, students may explicitly and implicitly attend to and coordinate any combination of these aspects and processes. Although the results are not conclusive, the differences in the two students' abilities to address the interviewer's probing questions suggests that students who attend to and coordinate more aspects and processes have constructed a more detailed and robust understanding of sampling distributions. This was seen in the responses by the other six students in the study.

These results suggest that the framework may be a useful tool for building on Saldanha and Thompson's frameworks to characterize in more detail the ways students think about repeated sampling and sampling distributions. In order to evaluate the framework further, it will be useful to conduct additional interviews in which the probing questions are designed to highlight the aspects and processes that appear to be either absent or only implicit to determine the degree to which explicitly coordinating all of the components of the framework is essential for understanding and working with sampling distributions.

References

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In J. B.-Z. Garfield, D. (Ed.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking.* (pp. 147-168). Springer.

Chance, B., DelMas, R., & Garfield, J. (2004). Reasoning About Sampling Distributions. In D. Ben-Zvi & J. Garfield, *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 295-323). Springer.

delMas, R.C., Garfield, J. and Chance, B.L.: 1999, *Exploring the role of computer simulations in developing understanding of sampling distributions,* Paper presented at the American Educational Research Association, Montreal.

Garfield, J., delMas, R., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM—The International Journal on Mathematics Education,* 44(7), 883-898.

Lipson, K. (2003). The role of the sampling distribution in understanding statistical inference. *Mathematics Education Research Journal*, *15*(3), 270-287.

Mills, J. D. (2002). Using computer simulation methods to teach statistics: a review of the literature. *Journal of Statistics Education* [Online], 10(1). www.amstat.org/publications/jse/v10n1/mills.html

Pfannkuch, M., & Reading, C. (2006). Reasoning about distribution: A complex process. *Statistics Education Research Journal, 5(2),* 4-9.

Saldanha, L. & Thompson, P. W. (2002a). Students' scheme-based understanding of sampling distributions and its relationship to statistical inference. In D. Mewborn (Ed.), *Proceedings of the Twenty-fourth Annual Meeting of the International Group for the Psychology of Mathematics Education*. Athens, GA.

Saldanha, L. & Thompson, P. W. (2002b). Conceptions of sample and their relationships to statistical inference. *Educational Studies in Mathematics, 51,* 257-270.

Saldanha, L., & Thompson, P. (2006). Investigating statistical unusualness in the context of resampling. *Proceedings of the International Congress on Teaching Statistics.* Bahia, Brazil.

Saldanha, L., & Thompson, P. (2007). Exploring connections between sampling distributions and statistical inference: An analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education.* (2), pp. 270-297.

Sedlmeier, P. (1999). *Improving Statistical Reasoning: Theoretical Models and Practical Implications*, Lawrence Erlbaum, Mahwah, NJ.

Strauss, A. & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage Publications.

Well, A.D., Pollatsek, A., and Boyce, S.J. (1990). Understanding the effects of sample size on the variability of the mean. *Journal of Organizational Behavior and Human Decision Processes*, 47, 289-312.

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., Chang, B., et al. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education*, *16*(2).

WHAT IS SIMPLIFYING?: USING WORD CLOUDS AS A RESEARCH TOOL

Benjamin Wescoatt
Valdosta State University

This paper describes the utilization of word clouds within a research methodology. To explore student notions of the concept of "simplify" in a trigonometry course, students responded to the prompt "In your own words, what does it mean to simplify?" The researcher created a word cloud derived from the student responses to explore and identify themes. These themes formed an initial framework for an in-depth analysis of the responses. During the textual analysis, the word cloud was consulted to confirm findings. Using the word cloud in preliminary and confirmatory roles adhered to the framework put forth by McNaught and Lam (2010). From the analysis, students appeared to view the act of simplifying as a process of taking an expression to its most basic state in order to reduce the perceived size (physical or cognitive) of the expression. Moreover, word clouds played a valuable role, providing visual representations of data.

Keywords: Word clouds, Simplify, Trigonometry, Data visualization

As part of a larger study, a need manifested to understand the meaning that students applied to the notion of *simplify*. A recently developed research tool provided a "quick and dirty" starting point from which to proceed. This paper outlines the utilization of Wordle.net to generate *word clouds*. While the results of the final analysis will be shared, this paper will also focus on outlining the methodology and appropriateness of word clouds as a tool for research in undergraduate mathematics education.

Word Clouds

A new method of data analysis, deriving from Web 2.0 internet sites such as blogs and social media sites, is the tag cloud, also referred to as a content cloud or a word cloud. In their original setting, word clouds were "visual presentations of a set of words, typically a set of 'tags' selected by some rationale, in which attributes of the text such as size, weight, or color are used to represent features of the associated terms" (Rivadeneira, Gruen, Muller & Millen, 2007, p. 995). Generally, word cloud formation depends on the frequency in which words are used in textual passages. The more frequently a word is used, the larger the word appears in the cloud picture, relative in size to the least frequently used word in the cloud.

Tags are human-generated keywords used to categorize information found on websites; they provide a summary of the content for users of the site, serving as tables of content. As an example, blogs use word clouds to summarize the content of posts, with the words in the cloud being keywords generated by the blogger. Tags for popular or frequently visited content on a website may appear with a bolder color or a larger font size in the word cloud. Additionally, within the word cloud, the individual tags can be hyperlinked to resources and content of the site labeled with the particular tag (Rivadeneira et al., 2007; Bateman, Gutwin, & Nacenta, 2008).

Rivadeneira et al. (2007) proposed a scheme for user tasks that word clouds support. Users can engage in *searching* for a specific concept on the site. However, a specific concept does not need to be known a priori; thus, word clouds allow users to engage in *browsing* for general content. While *gisting*, or impression forming, users are able to form a general impression of the content of the site, observing prevalent concepts while at the same time noting underlying themes. Finally, word clouds can provide users the ability for *matching*, or recognizing, sets of data. Thus, word clouds may provide identifying features unique to the content of the site; users may be able to narrow down what the underlying concept of the site is through exploration of the associated word cloud.

While word clouds originated with web-based applications, the categories of Rivadeneira et al. have broad appeal as the tasks associated with the categories are not necessarily bound to websites. For example, researchers may generate clouds to *search* for the presence of a specific idea or theme in textual data. On the other hand, if limited knowledge exists of a phenomenon and the research is of an exploratory type, researchers may *browse* the clouds and observe emerging themes. Clouds can also be used to form a quick impression or get the *gist* of the data. Finally, clouds may be used to distinguish and recognize the underlying phenomenon being observed. Anything that can be analyzed through a content analysis can be visualized with a word cloud (Cidell, 2010). Used as a research technique, word clouds can quickly show what the phenomenon being studied is about, leading to researchers forming generally impressions (Gottron, 2009). Additionally, word clouds can quickly reveal differences among ideas in selections of written or spoken texts through a visual inspection and comparison of the pictures, illustrating any emerging themes (McNaught & Lam, 2010; Williams, Parkes & Davies, 2013).

Research with Word Clouds

Despite the inherent potential, word clouds have been used sparingly in research thus far. Cidell (2010) proposed word clouds as a method for exploratory analysis in geographic information systems. The method consisted of generating word clouds using public meeting transcripts and eco-friendly building articles. For each data source, the word clouds were then visually mapped to represent the geographic location from which the transcripts or articles originated. Then, the clouds were analyzed for within and across themes; a location was explored for what mattered most to that region based upon prominence of words in the cloud and then certain words were explored across regions to compare the prominence. Results from the meeting transcripts were triangulated with comments made at the meetings; in this way, the word clouds served a confirmatory role, supporting the content analysis of the meeting transcripts. Cidell concluded that using the word clouds illuminated the differences in regional attitudes and commented that the word clouds suggested many avenues for future research into the issues being researched. Overall, Cidell maintained that the method of word clouds "offers the potential of combining content analysis, visualisation and qualitative GIS" (p. 522).

Williams, Parkes, and Davis (2013) used word clouds to gain an initial overview of aspects from an induction program in management education. Students responded to prompts exploring their views about what they enjoyed the least and most and about what they found useful in the induction program. The raw data were used to generate an initial word cloud. Finding the resultant cloud uninformative, the survey responses were classified as negative and positive. Then two word clouds were created from the responses in each category. Creating these thematic clouds allowed the data to be presented in a meaningful way, within their original context. Finally, common phrases were categorized into general themes, and word clouds were generated from these themes. The researchers concluded that the word clouds were powerful tools for preliminary research, allowing the data to be quickly analyzed.

Finally, McNaught and Lam (2010) discussed two studies in which word clouds served different roles. The first study explored human factors affecting the comments of participants in focus-group meetings. Before an in-depth analysis of the transcripts of the meetings, the transcripts (with minimal corrections) were used to create clouds. Cursory observations of the clouds led the researchers to note important differences among the meetings, providing a preliminary understanding of what occurred in the meetings and directing attention to issues that needed follow-up studies. In the second study, the researchers used comments from five participant blogs about the usage of eBooks to form five separate word clouds. Before inputting the text, researchers made minor modifications to the text to maintain a sense of context; this step was taken as the themes within the blogs were important to the researchers. For example, spaces after the word "not" were removed to retain a negative connotation of an idea. The word clouds were then compared in order to confirm previous analysis of students' perceptions of eBooks.

Word clouds do have their inherent limitations. Bateman et al. (2009) found that larger font size or larger font weight for words exerted the strongest influence on individuals exploring word clouds. Thus, larger or bolder words, indicating high prevalence in the text, could distract an observer from underlying themes, thus biasing the analysis. Additionally, Cidell (2010) pointed out that the sizes of words were relative to the frequencies of the other words in the text; the implication was that when comparing word clouds, similarly sized words do not imply that the word was mentioned the same number of times in the different passages. To circumvent this limitation and allow for better cloud comparison, researchers have suggested techniques such as parallel tag clouds (Collins, Viegas, & Wattenberg, 2009), seam carving (Wu, Provan, Wei, Liu, & Ma, 2011), and word storms (Castella & Sutton, 2013).

Also related to the usage of frequencies in forming the cloud image, McNaught and Lam (2010) suggested that word clouds should only be used to analyze the actual spoken (transcribed) or written word of participants. Using word clouds for field notes or researcher summaries would be less powerful as it would reveal information about the researcher and not necessarily the participants. Another drawback highlighted by McNaught and Lam was that word clouds remove the words out of their contexts. Thus, words being prominently displayed in a picture implied nothing about the importance of the word to the phenomena being investigated. Instead, prominence of a word merely suggested a further textual analysis into how the word was actually used in conjunction with ideas. In this way, the word cloud served a preliminary role, fostering the development of an initial thematic framework through which to conduct further analysis.

Methodological Framework

Therefore, due to the limitations of word clouds, McNaught and Lam (ibid) suggested that the role word clouds serve in research should be limited to a complimentary research tool. Specifically, they believed that word clouds are best used as:

- A tool for preliminary analysis, quickly highlighting main differences and possible points of interest, thus providing a direction for detailed analyses in following stages; and
- A validation tool to further confirm findings and interpretations of findings. The word clouds thus provide an additional support for other analytic tools. (p. 631)

The utilization of word clouds in this current study followed the McNaught and Lam framework.

Conceptual Framework

The premise for the investigation into students' conceptions of *simplification* was that the meaning of simplification held by a student was constructed on an individual level that developed within an influential classroom culture, a community of learners. That is, when comparing two mathematical expression, identifying which expression was "simpler" was not straightforward and a somewhat subjective task. For example, in comparing the two equivalent expressions $\sin 2x$ and $2 \sin x \cos x$, one student may state that $\sin 2x$ is simpler as it has a visually smaller physical size. On the other hand, a student uncomfortable with function

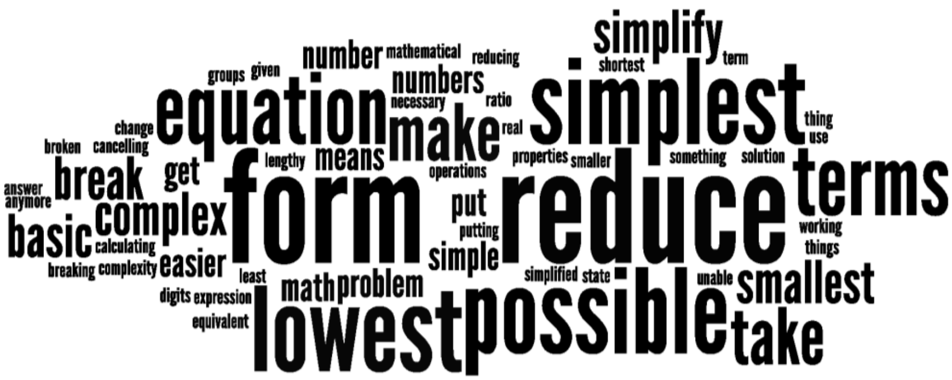


Figure 1. Initial word cloud of students' explanations of the meaning of "simplify."

arguments and double-angled trigonometric functions may be troubled by $\sin 2x$ and thus consider $2 \sin x \cos x$ to be the simpler expression. However, while meanings were individual, common conceptual facets existed across individuals due to the shared classroom culture.

Methodology

This study was conducted at a large, Midwestern research university within a college trigonometry course. In preparation for the first class period in the unit on verifying trigonometric identities, students responded to the prompt, "In your own words, what does it mean to simplify?" Of the 33 participants in the study, 24 participants completed the prompt. The participant responses to this question were analyzed in order to determine a consensus on the trigonometry classroom community's definition of what "simplify" meant.

The written responses were transcribed to a text document. Minimal grammatical and spelling corrections occurred; because the text was going to be used in an analysis that depended upon frequency, having a word spelled correctly was desired. This text in turn was copied and pasted to Wordle.net, and a word cloud (Figure 1) was generated. Wordle.net is an open-access website devoted to the quick formation of word clouds (dubbed "wordles"). At Wordle.net, the more frequently a word was used, the larger and bolder it appeared in the cloud.

After a brief scanning of the cloud, looking for prominence of words and similar ideas, a new text document was created from the original text document with the following changes. First, some students used the word "simplify" as a signal phrase in their responses, e.g., "To simplify, …." Thus, the word had little relevance for the explanations of what it meant and was edited out of the text document. Next, as "break" had some prominence, the similar phrases "breaking" and "be broken" were altered to "break" in order to capture the same idea. For the same reason, the word "reducing" was edited to "reduce." Finally, while enacting these changes, the observation was made that Wordle.net filtered the word "down" out of the cloud, treating it as a common *stop word*. However, students used "down" as a descriptive adverb in a significant way. Thus, a tilde was used to form the phrases "break~down" and "reduce~down," forcing Wordle.net to treat the phrases as a word and use them in the cloud formation. The new text document was fed to Wordle.net, and a new cloud (Figure 2) was generated.

A cursory analysis of the word cloud revealed several potential themes. First, the participants spoke of simplify in comparative states of existences, using words such as "lowest," "simplest," and "smallest." Size played a role in this existence, by virtue of the root word "small," "short," and "length." Actions such as "break down," "reduce," and "reduce down" were apparent in the cloud, suggesting a process. Along with these words, other prominent notions, such as "basic," "possible," and "complex" needed to be explored to identify the context in which they were used and to determine their relationships to the prominent words "form,"

Figure 2. Final word cloud of students' explanations of the meaning of "simplify."
"equation," and "terms." These observations suggested lines of inquiry for an in-depth textual analysis through the lens of this initial framework.

Findings

Having the notion of "to bring down or diminish to a smaller number, amount, quantity, extent, etc., or to a single thing; to bring down to a simpler form," ("Reduce", 2013), the word "reduce" is very prominent in the cloud. Thus, an analysis of the participant responses was undertaken to explore how students were using this idea as it related to simplifying. Students described "to simplify" as "to reduce" in eight distinct comments out of 24 total responses. Of these eight, one student directly equated simplifying with reducing, while another student explained simplifying as reducing the complexity. The other six comments were all related by the idea that simplifying was the act of reducing to some base state of existence with a notion of finality associated to it. For example, one student stated that to simplify meant:

- *Reduce problem to the lowest, simplest form possible.*

Thus, simplifying entailed changing the state of the expression to something that could no longer be changed; the primitive existence of the expression was reached.

The concept of reducing related to another prominent cloud feature, that of "break down." Although to break down may elicit several interpretations, students used it in the sense of decomposing something to simpler components. A directional quality of the act of simplifying was emphasized through pairing the word "break" with "down"; in the same manner of reducing, breaking down lead to a low-level, base state of existence for the expression. Once this simplest or basic form was reached, nothing further could be done. For example, students described what simplifying meant in the following way:

- *Breaking an equation down to the most basic form possible.*
- *To put in the simplest terms possible; an expression is unable to be broken down further once simplified.*

Thus, for these students, simplifying was akin to the atomism of Leucippus and Democritus; a point would be reached where the expression could not be further broken down and simplified.

While not specifically always using the phrases "reduce" or "break down" to describe the act, several comments described reaching a most primitive state, making the "form" of the expression more basic, simpler, or smaller. Examples of these comments were:

- *To get to the most basic form.*
- *Change it to the simplest form.*

In total, eight comments referred to the form of the expression dictating when the expression was simplified. Moreover, when writing of a "simplest" form, some students might have been thinking of the complexity in terms of the physical size of the expression, as suggested by the student who wrote:

- *To take a lengthy or complex equation and use mathematical operations to put the equation in its least complex form.*

In this comment, the student linked the length of the "equation" with a progression of the expression to its simplest form. (Note: We interpreted the meaning as "expression.")

Students also referred to simplifying as bringing the expression to its lowest or simplest "terms" (5 comments). One student explained simplifying in the following way:

- *Reducing to the lowest terms necessary.*

Colloquially, terms generally mean the components of something, so "lowest terms" might be taken to imply reaching a state that can be broken down no further. Also, some students might have believed simplest and lowest to be synonymous, as suggested by the following comment:

- *To break something down into its lowest/simplest terms.*

Furthermore, bringing an expression to lowest terms, to make it simpler, was connected to the physical size of the expression. As a student explained:

- *To make more simple by cancelling terms out.*

Cancelling terms out reduced the matter composing the expression, thus making the expression smaller in size.

The size of the expression was an underlying theme across many of the responses. In many instances, the size described the final state of the expression. However, as previously discussed, the notion of "making small" was embedded in the phrases "break down" and "reduce," as each phrase described simplifying as an action of taking the expression to a more basic state. The size of the expression may not necessarily refer to the physical size, but to a cognitive load. The more complex an expression was, the larger that expression was. Thus, simplifying shrank the cognitive load of the expression for that student. While the cognitive size might be related to factors based upon the actual components used and not depend on the size, e.g., a student may perceive $\sin x$ as being cognitively smaller than $\sec x$ due to familiarity with the sine function, students definitely linked complexity of an expression to the physical size.

Conclusions

To summarize, participants within the classroom community viewed the act of simplifying as a process of taking an expression to its most basic state in order to reduce the perceived size (physical or cognitive) of the expression. As some students wrote, simplifying meant:

- *To reduce the #'s that you are working down to the lowest possible digits to make the math easier.*
- *To take a complex thing and make it easier.*

This result was not groundbreaking or unexpected and aligned with a broader accepted meaning of simplify. However, establishing what students meant when using the word "simplify" was important for the study since simplifying acts played an important role in the larger issue being investigated, verifying trigonometric identities. For future study, a related issue is the role that context plays in students' perception of simplification. Will a form considered simple by an algebra student also be considered simple by a calculus student? Moreover, will the task involved and the needed representations for the task influence the perception?

Word clouds proved to be a valuable research tool in exploring the student notions. Since a community definition of "simplify" was desired, pre-existing frameworks to guide the exploration were shunned to bracket preconceived notions. Word clouds provided a quick entry point, allowing the development of an initial framework from the participant data through which to conduct the deeper analysis. Moreover, once the cloud was generated, it could be used as a touchstone during the in-depth analysis. In this way, it served to help verify the results of the textual analysis.

Another benefit of the word cloud was the quantitative visualization of the qualitative data. That is, word clouds have the ability to visually display qualitative data in a meaningful way. Themes evolving from the data may be readily apparent not only to the researcher but to the audience as well. Being able to visualize qualitative data has been an area of recent concern (e.g., Slone, 2009). Furthermore, the themes appear in a quantized way, depending on frequencies. Thus, word clouds should be considered as a viable option for the presentation of data in mixed methods research in addition to their utilization as a research tool.

References

Bateman, S., Gutwin, C., & Nacenta, M. (2008, June). Seeing things in the clouds: the effect of visual features on tag cloud selections. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia* (pp. 193-202).

Castella, Q., & Sutton, C. (2013). Word storms: Multiples of word clouds for visual comparison of documents. *arXiv preprint arXiv:1301.0503v1*.

Cidell, J. (2010). Content clouds as exploratory qualitative data analysis. *Area*,*42*(4), 514-523.

Collins, C., Viegas, F. B., & Wattenberg, M. (2009). Parallel tag clouds to explore and analyze faceted text corpora. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on Visual Analytics Science and Technology* (pp. 91-98).

Gottron, T. (2009). Document word clouds: Visualising web documents as tag clouds to aid users in relevance decisions. In *Research and Advanced Technology for Digital Libraries* (pp. 94-105).

McNaught, C., & Lam, P. (2010). Using Wordle as a supplementary research tool. *The Qualitative Report, 15*(3), 630-643.

Reduce, v. Def. I.2a. (2013). In *OED Online*. Oxford University Press. Retrieved from http://www.oed.com/view/Entry/160503?rskey=mY1LVm&result=2&isAdvanced=false#eid

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007, April). Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 995-998).

Slone, D. (2009). Visualizing qualitative information. *The Qualitative Report, 14*(3), 489-497.

Williams, W., Parkes, E. L., & Davies, P. (2013). Wordle: A method for analysing MBA student induction experience. *The International Journal of Management Education, 11*, 44-53.

Wu, Y., Provan, T., Wei, F., Liu, S., & Ma, K. L. (2011). Semantic-preserving word clouds by seam carving. *Computer Graphics Forum, 30*(3), 741-750.

STUDENT CONCEPTIONS OF TRIGONOMETRIC IDENTITIES THROUGH APOS THEORY

Benjamin Wescoatt
Valdosta State University

This preliminary study attempts to describe an initial genetic decomposition of a trigonometric identity for college students. Scant research exists into the concepts found in trigonometry. Thus, little is known about how students actually understand a trigonometric identity. Following the guidelines of APOS theory, an initial genetic decomposition for a trigonometric identity was proposed. According to this decomposition, students with action conceptions can verify identities explicitly using step-by-step manipulations while students holding a process conception are able to visualize steps to demonstrate that the identity is true. Having an object conception means students recognize the truth of the equality without verification and are able to then use the identity to verify other identities. After observing students in task-based interviews, needed modifications to the genetic decomposition became apparent. For example, students' conceptions of the function argument appeared to influence the verification process.

Key words: Trigonometric identity, APOS theory, Function, Equal sign

A trigonometric identity is an object encountered by many high school and college students. The identity itself is a tautological statement claiming that two expressions composed of certain combinations of trigonometric functions actually describe the same underlying mathematical object despite appearing to be different. While the notion of the trigonometric identity is rich in mathematical concepts, to date, research into students' notions of the trigonometric identity is scarce. Moreover, how students' understandings of the identity develop is virtually non-existent.

APOS Theory

Rooted in the theories of Piaget, APOS theory attempts to describe how students may come to understand certain mathematical objects. Underpinning APOS is the hypothesis that

> an individual's mathematical knowledge is her or his tendency to respond to perceived mathematical problem situations and their solutions by reflecting on them in a social context and constructing or reconstructing mathematical actions, processes and objects and organizing these in schemas to use in dealing with the situations. (Dubinsky, 2000, p. 11)

According to Dubinsky and McDonald (2001), an action is a learner-perceived external transformation of an object. Actions usually occur in a step-by-step fashion with reliance on memorized procedures. Once the learner has repeated an action, reflection on the action may interiorize the action into a process. A process does not need to be physically performed; the learner may envision the process and the result. Thus, a process does not have a reliance on external stimuli but is under the control of the learner. Once the learner understands that a process represents a totality, the learner is said to have encapsulated the process into an object. The learner constructs the mathematical concept's schema by collecting into a coherent framework all of the other actions, processes, objects, and schemas associated with that concept.

The schema is composed of a genetic decomposition in addition to the linked concepts. The genetic decomposition of a mathematical concept is a proposed model of cognition, or "a structured set of mental constructs which might describe how the concept can develop in the mind of an individual" (Asiala et al., 1996, p. 13). Thus, the genetic decomposition attempts a reasonable explanation for how a learner might come to understand the concept and is based upon observable phenomena. In order to develop the decomposition, a researcher begins with a learning theory and previous observations and constructs a hypothesis. This hypothesis serves as the initial lens through which to interpret data. Analysis of the data refines the hypothesis, and the decomposition evolves to better explain the observed phenomena.

Problem Statement

As little is known about trigonometric concepts, this study aims to fill the void in the literature related to identities. Specifically, this study will propose to describe college students' conceptions of trigonometric identities through the following questions.

1. What is the genetic decomposition of a trigonometric identity?
2. In what other ways do students encapsulate trigonometric identities into objects?

The last question is for the purpose of exploring the appropriateness of using APOS theory with trigonometric identities; as suggested by Tall, APOS theory should not be taken as a global theory of learning (Tall, 1999).

Related Literature

Breidenbach, Dubinsky, Hawks, and Nichols (1992) explored students' conceptions of function through APOS theory. From an analysis of student survey responses, they revised their initial genetic decomposition. Students with action conceptions of function needed an explicit formula or recipe to follow. The recipe would then be followed in a step-by-step fashion. The emphasis in an action conception was on the external recipe. Students could not rely on any relationships existing in their mind. On the other hand, students with a process conception of function generally viewed the function in an internal way. That is, the function did not need to be given explicitly for the student to think about the transformations of the function. This aspect of the process conception is especially cogent for trigonometry as students traditionally encounter the trigonometric functions sans the explicit formula (Weber (2005) demonstrated how students could successfully experience trigonometric functions in a prescriptive sense).

A schema for the trigonometric identity would extend beyond function. Thus, a learner's schema for function would be treated as an object as part of the identity schema. Clark, et al. (1997) described how the function schema became thematized to processes and objects within the schema of the chain rule. After their analysis of students' conceptions of chain rule through their initial genetic decomposition, they introduced the Piagetian Triad (Piaget & Garcia, 1989) in order to better explain the development of the chain rule schema.

Trigonometric identities represent a special equality. Research into conceptions of the equal sign generally describes students' conceptions of the equal sign in terms of *operational* versus *relational* (Kieran, 1981). In terms of a hierarchy, the operational meaning develops first before the relational meaning manifests in students' understanding (McNeil & Alibali, 2005). However, studies have found that students of all ages, even within the college ranks, may hold a weak notion of the equal sign, viewing it operationally rather than relationally (Kieran, 1981; Weinberg, 2010).

Genetic Decomposition

The manipulations of identities are taken to be the verifications of the identity, where being verified is taken to be relative to the learner. As part of the identity schema, students should hold a process or object conception of function. With an action conception of a trigonometric identity, students verify the identity by explicitly writing out steps. However, with a process conception of trigonometric identity, students visualize several steps of the process and "see" or "feel" the path to take. With both the action and process conceptions, students cannot use the identity to verify another identity prior to verification. Students with an object conception of the trigonometric identity accept the identity as being true without having to verify it. If needed, they are able to de-encapsulate the identity into its string of equalities. The schema for a trigonometric identity consists of all identities being used to build knowledge in various settings.

Methodology

The data for this study was collected from a college trigonometry course at a large research university as part of a larger study on verifying trigonometric identities. Thirty-three students participated, responding to prompts involving verifying identities and solving verification problems. Of these thirty-three students, eight agreed to participate in individual task-based interviews. Each interviewee solved verification problems while speaking aloud his or her thought processes. The audio from the interviews was captured and transcribed.

While reading each interview, general themes relating to the initial decomposition were noted and compared to the previously-read interview transcript. These themes were triangulated within and across interviews in order to test the explanatory power of the initial decomposition. This analysis process is currently ongoing and will cycle until a satisfactory evolved genetic decomposition is formed.

Initial Findings

Although analysis is ongoing, some findings are emerging, indicating needed modifications to the genetic decomposition. First, the role of the function argument was not accounted for in the decomposition. Some students indicated a preference for working with only the variable *x* while others showed fluidity in working with complex arguments. For example, while solving a problem, Katie accidently wrote *x* instead of *y* and explained, "I always have *x* in my head, so I'm bad about writing *x*. … I would rather just write *x* just cuz it's natural for me to write *x*." On one problem, students spoke of seeing an "*x*" in their head even though the function argument was $2\alpha - 1$. Function arguments that were not *x* presented barriers for students, as Cooper described, "Whenever I don't see *x*, I kind of ignore the identities for a moment until I look at it and go, oh, it's the same thing. It's the same thing as saying cosine two *x* or cosine two *y*." Because students expressed differing comfort levels with the function argument and because it presented barriers to students, the function argument should be included in the genetic decomposition. This makes sense as the truth of an identity depends on the truth of the equality for all argument inputs. As Amber stated, "You're not actually verifying that tangent *y* plus cotangent *y* over cosecant *y* times secant *y*, um, is equal to one because of *y*. You're not thinking it's because of that variable."

The data appeared to support encapsulation. When asked whether an equation was an identity, Cooper stated that he believed it was not, but rather, it was a definition. For Cooper, identities represented equations that had been verified while definitions were identities that had been encapsulated to an intimate level of familiarity. He was able to immediately recall definitions to mind but paused with identities. This pause may have been due to a need to verify, in his mind, that the identity was true. As he stated in describing the hesitation, "Yeah, to kind of get an idea, and try to make sure you got it right. Kind of roll through the process some." On the other hand, definitions required no work. In describing the Pythagorean identity, a definition for him, he claimed, "It doesn't really require you to change it up any at all. It's just kind of defining what one is." Thus, Cooper appeared to have encapsulated some identities as objects.

The difference between action and process conceptions as being external versus internal was supported by a metaphor some students used to describe their work, an "unraveling." Unraveling started a process in which correct steps to take become readily apparent to the student, guiding decisions. As Alan described, "Once I got a pretty good starting point, that's when I just went and, as I was doing it, it kind of unraveled itself a bit." Students reported of searching for paths to take consisting of multiple equivalent expressions, while other students appeared to take each step individually before proceeding to the next step.

Conclusions and Audience Questions

The initial genetic decomposition needs to evolve to better account for the observed phenomena. While the full analysis is not finished, APOS theory appears to be able to explain much of the process of verifying trigonometric identities. One noted shortcoming is that the data comes from algebraic verification attempts. Whether or not APOS theory is adequate to describe geometric verifications would need to be explored.

1. Tall, Thomas, Davis, Gray, and Simpson (2000) explored what the object of the encapsulation was across several theories. Would it be more appropriate to view a trigonometric identity through another lens, such as a procept?
2. Equality should play a role. How could the categories of *operational* and *relational* be integrated into the genetic decomposition?

References

Asiala, M., Brown, A., DeVries, D., Dubinsky, E., Mathews, D., & Thomas, K. (1996). A framework for research and curriculum development in undergraduate mathematics education. *Research in Collegiate Mathematics Education II,* 1-32.

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics, 23*(3), 247-285.

Clark, J., Cordero, P., Cottrill, J., Czarnocha, B., DeVries, D. J., St. John, D., … Vidakovic, D. (1997). Constructing a schema: The case of the chain rule? *Journal of Mathematical Behavior, 16*(4), 345-364.

Dubinsky, E. (2000). Using a theory of learning in college mathematics courses. *MSOR Connections, 1*(2), 10-15.

Dubinsky, E., & McDonald, M. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton (Ed.), *The Teaching and Learning of Mathematics at University Level: An ICMI Study* (pp. 275-282). Dordrecht: Kluwer Academic Publishers.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*(3), 317-326.

McNeil, N. M., & Alibali, M. W. (2005). Knowledge change as a function of mathematics experience: All context are not created equal. *Journal of Cognition and Development, 6*(2), 285-306.

Piaget, J., & Garcia, R. (1989). *Psychogenesis and the history of science*. (H. Feider, Trans.). New York, NY: Columbia University Press. (Original work published 1983)

Tall, D. (1999). Reflections on APOS theory in elementary and advanced mathematical thinking. In O. Zaslavsky (Ed.), *Proceddings of the 23rd Conference of PME, Haifa, Israel,* 1, 111-118.

Tall, D., Thomas, M., Davis, G., Gray, E., & Simpson, A. (2000). What is the object of the encapsulation of a process? *Journal of Mathematical Behavior, 18*(2), 223-241.

Weber, K. (2005). Students' understanding of trigonometric functions. *Mathematics Education Research Journal, 17*(3), 91-112.

Weinberg, A. (2010). Undergraduate students' interpretations of the equals sign. In *Proceedings* *of the 13th Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC.

Developing Pre-Service Secondary Math Teachers Capacity with Error Analysis Related to Middle-Grades Mathematics

As part of a National Science Foundation Noyce Scholarship Grant, one university substantially revised its preservice secondary (grades 7-12) math teacher preparation program. As one component of this program, preservice teachers take three credit hours of middle level number and operation and geometry, with a focus on mathematical knowledge for teaching. As a research component, we investigated the impact of this course on preservice teachers capacity to identify, analyze, and respond to student errors. This paper provides additional background and results from the first two offerings of the course, as well as ideas for further study.

Key words: Error Analysis, Mathematical Knowledge for Teaching, Number and Operation, Geometry, Preservice Secondary Teachers

Background and Perspectives

The supply of qualified, competent mathematics majors entering secondary teaching professions is not keeping pace with demand (Liu et al., 2008; National Research Council, 2002). According to Ingersoll and Perda (2009), the problem is more than a number game. Part of the problem resides in the fact that teachers are not happy with the profession once they are out in the fields, causing the number of teaching leaving the profession to be greater than the number of teachers entering the profession. This is especially the case in low-income areas where they are 77% more likely to be taught by out of field teachers compared to students from high socioeconomic backgrounds (Ingersoll, 2003). Attrition is a compounding factor as recent research reveals 20-30% of teachers have left the profession within the first five years (Darling-Hammond, 2001).

Specifically, this preliminary report discusses how the Rocky Mountain Noyce Scholars Program, a five-year National Science Foundation scholarship grant for undergraduate pre-service secondary mathematics teachers, aims to target undergraduate education as part of the solution. The idea is that if we recruit teachers who are strong in their content area and also dedicated to serving students in high needs school districts that the attrition rates may decrease. If we also help prepare and support our teachers well (both in the areas of mathematics and pedagogy), they will hopefully be both successful and interested in remaining in the teaching profession.

Adhering to these ideals, the Rocky Mountain Noyce Scholars Program has been a catalyst for revision of the undergraduate secondary mathematics teacher preparation program. We discuss one component of this revision in detail in conjunction with preliminary results from the first two years of the program. In our presentation, we will seek advice from our audience members on future research steps and data collection to strengthen the preliminary results and contribute to national knowledge related to best practices in preparing teachers.

Program, Participants, and Context

In the author's state, there is no separate middle level math certification for teachers. As such, secondary teacher preparation programs are designed to prepare teachers for state level certification for grades 7-12. Most of these (certainly all that the author is aware of) focus primarily on the high school grades, implicitly assuming that this prepares teachers to teach middle level mathematics.

However, it is by now widely accepted in the mathematics education community that there is specialize mathematical knowledge for teaching (Ball, Thames & Phelps, 2008; Shulman, 1986). Before the revision, our program focused on preparing the teachers through a traditional mathematics major, with no attention to any specialized mathematical preparation.

Also, recommendations from the Conference Board of the Mathematical Sciences Mathematical Education of Teachers II document note that

> *Because the middle grades are 'in the middle', it is critical that middle grades teachers be aware of the mathematics that students will study before and after the middle grades.*

Additionally, the document notes that middle-grade teachers'

> *perspective on what it means to know mathematics may be based on their own success in learning facts and procedures rather than on understanding the underlying concepts upon which the procedures are based.*

To address these recommendations and observations, our revision added seven credit hours to the mathematics major aimed specifically at the mathematical knowledge needed for teaching mathematics. One component of this revision is three credit hours of content in number and operation and geometry. This course focuses on mathematical knowledge for teaching at the middle-school level, specifically on the mathematics needed for teaching in grades 4-8. The formal catalog description states:

> *Advanced study of number and operation, including why the various procedures from arithmetic work and connections to algebraic reasoning. Focuses on using rigorous mathematical reasoning and multiple representations to explain concepts.*

We aim to exposure the preservice teachers to common student misconceptions, elementary and middle level concepts related to number and operation as well as geometry, and in general the associated specialized knowledge for teaching these topics. The primary texts for the course were those by Beckmann (2008) and Ma (1999).

While we collected a variety of data for assessment, evaluation, and research purposes, the primary research question associated with this preliminary report is the following: Describe how a specialized course in number and operation impacts pre-service teacher ability to identify, analyze, and mathematically respond to common student errors.

Data Collection and Data Analysis

Data was collected during the first and last week of the course, during two consecutive offerings of the course. Specifically, the pre-service teachers were given an instrument consisting of four items to analyze and respond to. Each item dealt with a common student misperception at the upper elementary level (grades 4-8), and began by showing several examples of student work that were incorrect. The pre-service teachers were then given several example problems and asked to solve them in the same way that the student did. After that, they were asked to respond to the following prompts, based heavily on those from Ashlock (2009):

(a) Describe the procedure that this student seems to be using.

(b) What does this student seem to understand about (e.g., multiplying fractions)?

(c) What does this student not yet seem to understand about (e.g., multiplying fractions)?

(d) Has this student learned another concept or procedure that s/he is confusing with this one?

(e) Describe in detail multiple ways that you could help this student or other students with a similar misconception?

Preliminary Results

A qualitative and quantitative analysis of pre-service teacher responses shows several outcomes:

(a) a higher percentage of pre-service teachers identified the error at the end of the course than at the beginning,

(b) pre-service teacher were able to provide a much more in-depth analysis of what the student may and may not yet understand,

(c) pre-service teachers were able to both provide more potential ways to help the students, as well as to describe these ways in considerably more depth, and

(d) at the beginning of the course, pre-service teacher responses were predominantly procedural and algorithmically oriented, whereas at the end of the course, they were much more conceptually oriented.

Discussion

We will seek input from the audience on how to best research the various impacts of this course on students. A large part of the desire to give this talk is to solicit audience input on how to more methodically pursue our investigation. The author is a mathematician who is in the process of learning how to conduct mathematics education research. Feedback and input from this group would be particularly welcome. Some questions that arise for discussion include:

1. How can we better analyze impacts of this course on preservice secondary teachers?
2. Is such a course necessary for preservice secondary math teachers (certified 7-12), or are these skills and knowledge developed with sufficient depth in other ways?
3. Does this increase in skill related to error analysis transfer to other domains (e.g., to error analysis of, say, algebra)? If so, how could this be measured?
4. How might we isolate the components of this course that are central to the increase in pre-service teacher gains in error analysis?

References

Ashlock, R.B. (2009). *Error Patters in Computation: Using Error Patterns to Help Each Student Learn* (9th ed.). Pearson/Addison Wesley.

Ball, D. L., Thames, M. H., & Phelps, G., (2008). Content knowledge for teaching: what makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Beckmann, S. (2007). *Mathematics for elementary school teachers* (2nd ed.). Boston, MA: Pearson/Addison Wesley.

Conference Board of the Mathematical Sciences (2012). The Mathematical Education of Teachers II. Providence RI and Washington DC: American Mathematical Society and Mathematical Association of America.

Darling-Hammond, L. (2001). The challenge our schools face. *Educational Leadership, 58,* 12-17.

Ingersoll, R. M., & Perda, D. (2009). The mathematics and science teacher shortage: Facts and myth. Available at http://www.cpre.org/images/stories/cpre_pdfs/math%20science%20shortage%20paper%20march%202009%20final.pdf

Ingersoll, R. (2003). *Is there really a teacher shortage?* Philadelphia, PA, and Seattle, WA: Consortium for Policy Research in Education, University of Pennsylvania, and the Center for the Study of Teaching and Policy, University of Washington.

Liu, E., Rosenstein, J., Swann, A., Khalil, D. 2008. When Districts Encounter Teacher Shortages? The Challenges of Recruiting and Retaining Math Teachers in Urban Districts. *Leadership and Policy in Schools.* 7(3): 296-323.

Ma, L. (1999). Knowing and teaching elementary mathematics: Teachers understanding of fundamental mathematics in China and the United States. Mahwah, NJ: Lawrence Erlbaum.

National Research Council. (2002). *Learning and understanding: Improving advanced study of mathematics and science in U.S. schools*. Washington, DC: National Academies Press.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

CHARACTERIZING MATHEMATICAL COMPLEXITY OF TASKS IN CALCULUS I

Nina White, University of Michigan
Vilma Mesa, University of Michigan
Cameron Blum, University of Michigan

We present findings from a revised framework created to analyze tasks that calculus teachers assign their students. In the presentation we will highlight the features of the analytical framework and the steps taken to ensure high inter-coder reliability. The framework has been used to analyze all tasks (N=2,996) present in homework, quizzes, and exams from six faculty teaching Calculus I in two two-year colleges. We highlight some insights we have gained in creating this framework and possible uses by other researchers and other contexts.

Key words: Task analysis, Graded work, Ungraded work, Cognitive demand, Representations of functions, Calculus, Two-year colleges,

The complexity of mathematical tasks is frequently used to determine the quality of instruction. Indeed, instruction that reduces the complexity of the mathematical tasks is known to be detrimental for students' performance in standardized tests in middle school mathematics (Silver, Smith, & Nelson, 1995; Silver & Stein, 1996). Several frameworks that determine this complexity have been proposed. We document the development of one framework that allows us to assess this complexity for exams, quizzes, homework, and generally for any type of graded and ungraded work that teachers assign in Calculus I courses. We discuss the challenges and successes in creating this framework as well as its anticipated uses.

Theoretical Framework

This work rests on four dimensions that have been used to establish the complexity of mathematical tasks assigned to students. First is the type of knowledge that the task elicits (factual, procedural, conceptual, or metacognitive) together with the type of cognitive processes that could be hypothetically involved. The cognitive demand of a problem is a major determinant of task complexity in the literature (Anderson et al., 2001; Silver, Mesa, Morris, Star, & Benken, 2009). The second dimension, derived from the literature in mathematics education refers to the types of representations that are called for in the task, with agreement that translations across representations (e.g., graph to symbolic or verbal) are more demanding that transformations within a representation (Janvier, 1987; Kaput, 1992). The third dimension refers to argumentation in mathematics. As justification is a central part of mathematics, problems that require students to provide justification of their claims can be seen as mathematically more complex. The fourth dimension refers to opportunities to model with mathematics. The ability to "mathematize" a situation is one of the major skills we hope students in Calculus I to learn.

We attend to these dimensions in the task analysis because if the learning goals of Calculus I instruction include novel reasoning, fluent translation between representations, and skills in modeling and justification, then the tasks given to students should display those characteristics, so students can indeed practice them and become proficient in their use.

Method

Methodological work on this project builds directly on a framework developed by Tallman and colleagues (Tallman, Carlson, Bressoud, & Pearson, 2012) who analyzed 3,735 tasks present in exams from 150 institutions participating in the national study of Calculus I (Bressoud, Carlson, Mesa, & Rasmussen, 2013).

The data used to refine this framework consisted of 2,996 Calculus I tasks coming from the complete set of quizzes, exams, homework (both graded and ungraded) from six instructors at two southern community colleges. These tasks were collected as part of the case studies conducted by the project on the *Characteristics of Successful Programs in College Calculus* (CSPCC).

Table 1: Definition of Codes in each Dimension.

Item Orientation[a]	
Procedure	Direct prompt to use a specific method or procedure.
Recognize & use Procedure	Recognize and apply procedure. This covers cases where procedure is not given, but where the ultimate task is procedural. Students must recognize what procedure to apply.
Understand	Make interpretations, make comparisons or make inferences that require an understanding of a mathematics concept.
Apply Understanding	Use understanding to solve a problem when the method to be used is not directly proposed.
Item Representation, Prompt	
Numeric	Information crucial to solving problem conveyed in discrete form.
Symbolic	Information crucial to solving problem conveyed in symbols.
Graphical	Information crucial to solving problem conveyed graphically.
Verbal	Crucial information, other than basic instructions, conveyed in words.
Definition	Mathematical definition or theorem is provided for students to use in an axiomatic way, that is to apply or build off of
Theorem	Theorem-like statement is provided for students to consider validity of
Item Representation, Response	
Numeric	Students create or present novel information in a table or discrete form.
Symbolic	Students carry out symbolic manipulations or present/create novel information in symbols.
Graphical	Students create/present novel information in graphical form
Verbal	Students create/present novel information in words.
Proof	The task asks students to create a (fairly formal) proof.
Example	The task asks students produce an example or counterexample.
Extra Features	
Modeling	The task requires students to define relationships between quantities. The task may also prompt students to define or use a mathematical model to describe information about a physical or contextual situation.
Justification	The problem expects justification, in the form of computations or deductive logic, to describe *why* a claim is correct.

Note: a. The codes in the Orientation category are mutually exclusive, but the rest are not. We don't include codes that showed up in less than 1% of the coded tasks.

The first and third author worked in tandem to refine the definitions from Tallman's original framework to fit our goals and allow for reliable coding. Table 1 shows the four dimensions of the framework, with the codes used and a short definition for each. The coders worked through "coding sessions" in which they compared and discussed jointly-coded subsets of the data (as well as non-corpus data early in the process). Agreement in each dimension was computed using Cohen's Kappa. This was followed by discussion between coders guided by the agreement statistics, allowing us to refine the definitions and calibrate the coding process. Eventually, the third author coded all the tasks and the first author performed random checking of an additional 10% of the tasks. In the dataset as a whole, the two coders reached a Cohen's Kappa of 0.7 in all but three categories; such a Kappa is considered to be very strong agreement (Landis & Koch, 1977). See Table 2 for agreement on the final dataset. The difficulty in reaching agreement in Item Orientation is not surprising; judging the *potential* cognitive demand of a problem in the absence of student work or student thinking is notoriously difficult. The agreement in this dimension early in the process was much worse (as low as 0.14), and reaching this eventual agreement was a large part of the calibration effort.

Table 2: Cohen's Kappa for each dimension between the first and the third author on a randomly selected set of tasks (n=557) coded by the first author.

dim	**Orientation**	Format	Representation Prompt						
			G	V	S	N	Th	**D**	**mult**[a]
Kappa	**.47**	.75	.83	.76	.90	.76	.81	**.57**	**.54**
dim	J	AM	Representation Response						
			G	V	S	N	Pf	mult[a]	none[b]
Kappa	.65	.80	.77	.70	.80	.74	.85	.70	.76

Notes: a. "mult" refers to whether coders detected multiple representations. b. some questions do not require a written response and hence have no representation in their response.

Findings

The major accomplishment at this point is the calibration of the coding system; a full analysis of the coding has not been completed. Preliminary evidence confirms the results from Tallman and his colleagues, that tasks are predominantly procedural and symbolic (even though the tasks in this study come from institutions identified by CSPCC as being "successful"). However, there are certainly nuances to this breakdown that can be seen between types of tasks (e.g. homework vs. exams) and between institutions and instructors.

Discussion

As mentioned above, we have observed interesting differences (not yet tested statistically) regarding the different types of assignments. For example, graded homework tends to have relatively larger proportion of complex problems than ungraded homework. (Here we use "complex" to mean higher cognitive demand, more diverse representations, and more opportunities to justify and model.) This is not surprising, as ungraded homework tends to focus on procedural skill-building. More surprisingly, at one of the institutions, problems on exams tend to be more complex and difficult than those on homework. Perhaps in this case the instructor is not giving sufficient opportunities before the exam for students to use higher order cognitive functions. However, at another site we see an opposite phenomenon; we see more cognitively demanding problems on the homework than on exams. This is perhaps because it is seen as more fair to give harder problems outside of the timed exam setting.

We have also started to observe interesting variations by specific instructors (data not shown here, but to be included in the longer paper of this work). For example, at an institution where some faculty were present for a large calculus reform effort in the early 90s and some were not, those around for the original reform continue to use a much greater variety of representations in the tasks they assign. More broadly, what is starting to be clear from the data is that teachers have *signatures* regarding the graded and ungraded work they give their students. The notion of *signature* is borrowed from Stigler and Hiebert (1999) who in *The Teaching Gap* talked about the different teaching signatures present by country. In a setting like the two-year college it is not surprising that these signatures exists, given the greater autonomy of teachers relative to K-12 counterparts. Our current work on this project seeks to quantify that signature.

This framework will be used to code the rest of the task corpus of the CSPCC study and we anticipate interesting statistical differences between instructors, institution type, and task type (e.g. homework vs. exam).

Questions

The following questions will be presented to the audience for feedback:

1. In addition to coding the CSPCC task corpus, what other uses might this framework have?
2. How could such a framework be adapted to post-secondary mathematical contexts outside of Calculus I?

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., . . . Wittrock, M. C. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing*. New York, NY: Longman.

Bressoud, D. M., Carlson, M., Mesa, V., & Rasmussen, C. L. (2013). The calculus student: Insights from the MAA national study. *International Journal of Mathematical Education in Science and Technology*. doi: 10.1080/0020739X.2013.798874

Janvier, C. (Ed.). (1987). *Problems of representation in the teaching and learning of mathematics*. Hillsdale, NJ: Erlbaum.

Kaput, J. (1992). Technology and mathematics education. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 515-556). New York: Macmillan.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Silver, E. A., Mesa, V., Morris, K., Star, J., & Benken, B. (2009). Teaching for understanding: An analysis of mathematics lessons submitted by teachers seeking NBPTS certification. *American Educational Research Journal, 46*, 501-531.

Silver, E. A., Smith, M. S., & Nelson, B. S. (1995). The QUASAR project: Equity concerns meeting mathematics education reform in the middle school. In W. Secada, E. Fenemma & L. B. Adajian (Eds.), *New directions for equity in mathematics education* (pp. 9-56). New York: Cambridge University Press.

Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education 30*, 476-521.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap*. New York: The Free Press.

Tallman, M. A., Carlson, M., Bressoud, D. M., & Pearson, J. M. (2012). *A characterization of calculus I final exams in US colleges and universities*. [manuscript under review]. Arizona State University, Tempeh AZ.

ASSESSMENT IN UNDERGRADUATE INQUIRY-BASED LEARNING MATHEMATICS COURSES

Preliminary Research Report

Tim Whittemore — University of Michigan
Vilma Mesa — University of Michigan

We report initial findings of a study that seeks to investigate the methods instructors' use to assess their students' learning and how these assessments affect the instruction in their classrooms. Using data collected from 23 instructors using inquiry-based learning methods, we seek to discuss the instructors' goals for the students, the ways they measure the students' progress towards these goals, the feedback they give students, and how these assessments affected their instruction. Our analysis of the data uses open coding of the transcripts and of the documents (e.g., syllabi, exams, homework assignments) that the instructors gave to the students. Instructors cite using informal assessments and focusing on presentations when asked about "knowing" that students are learning. They cite formal assessments and examinations when asked about "measuring" that students are learning. We seek input on the analysis of the materials as current results may depend on the coding system used.

Keywords: Inquiry-based learning, Assessment, Instruction, Feedback

With this study we seek to fill a gap in the knowledge that exists about how instructors assess their students' learning in post-secondary mathematics courses. In particular, we look at the assessment methods reported by instructors teaching a range of inquiry-based learning [IBL] mathematics courses. Instructors, students, parents, administrators, and politicians often question what a student is learning in a class and question how instruction is helping this process. Assessment methods, defined as "the development of an awareness, through diagnosis, of [students'] progress" (Noss, Goldstein, & Hoyles, 1989, p. 109), are used to answer these questions. These methods can be used on a daily basis or at the end of the semester (Webb, 1995). The method of assessment defines "in measurable terms what instructors should teach and students should learn" (Stull, Varnum, Ducette, Schiller, & Bernaki, 2011, p. 30). Assessment methods can be used for many purposes. Diagnostic assessment can be used to identify knowledge students have mastered before learning activity. Summative assessment can be thought of as assessment *of* learning and formative can be thought of as assessment *for* learning. A fourth purpose, assessment *as* learning, is also possible (Torrance, 2007). In this purpose, assessment is on-going and the results help to shape the instructor's actions and the students learning opportunities (Torrance, 2007). The process of reshaping instruction, however, may be postponed to a subsequent semester (Davis & McGowen, 2007). Though much research about assessment in mathematics has been conducted in K-12 settings, less has been done in post-secondary mathematics, though there is a "plurality of assessment in university mathematics" (Iannone & Simpson, 2011, p. 186). In these settings, mathematical assessment provides a "comprehensive accounting of an individual's or group's functioning within mathematics or in the application of mathematics" (Webb, 1995, pp. 662-663). Part of the ongoing analysis will be to see how the assessment methods used by these instructors align with the material and applications emphasized in the classes.

The Supporting Assessment in Undergraduate Mathematics [SAUM] project examined the "tensions and tethers" that postsecondary mathematics instructors faced and how these limited the types of assessments that instructors were inclined or able to adopt (Madison, 2006). Madison notes that "…many instructional programs are tied to traditional in-course testing and have no plans to change, placing significant limits on assessment" (Madison, 2006, p. 4). This use of "traditional in-course testing" also includes methods such as extensive homework and final exams. Though the SAUM project focused on undergraduate mathematics, and some of the cases examined oral presentations and communication, IBL courses were not a primary focus of the study. Little research on student assessment has been conducted in IBL courses (Kogan & Laursen, 2012; Laursen & Hassi, 2010).

In IBL courses, there is an emphasis on students taking responsibility for their learning and presenting their work to their classmates (Maaβ & Artigue, 2013; Yoshinobu & Jones, 2012). The teaching methods used in IBL courses descends in part from the teaching style of R. L. Moore at the University of Texas. Moore believed students should build their own understanding and work through a pre-established sequence of problems and theorems. Focusing on and requiring students to present their work has shifted the teaching of these courses away from the lecture-based format often found in post-secondary mathematics settings (Coppin, Mahavier, May, & Parker, 2009). This shift in teaching may lead to a shift in the methods used to assess students. Recognizing this potential shift and an initial awareness of instructors' comments on teaching led us to ask three questions: how do instructors assess their students when teaching an IBL course?; how do the instructors' goals align with the assessment methods they report using?; and how do instructors' methods of assessment differ with regard to instructors' familiarity with IBL methods and the type of course they teach (e.g., lower division, upper division, courses for future teachers)?

Methods

Participants in this study were recruited from the 2011 and 2012 R.L. Moore Legacy Conferences. The courses these instructors taught were categorized as lower division (e.g., Calculus, Introduction to Proof), upper division (e.g., Real Analysis, Abstract Algebra), or courses for future teachers (e.g., Math for Elementary Teachers, Problem Solving for Prospective Secondary Teacher). In addition to these categories, we asked instructors to rate their familiarity with inquiry-based learning, as beginner, novice, advanced, or expert. We selected a sample of 23 instructors to interview. These instructors taught a range of courses and reported having a range of levels of familiarity. With this sample, we are able to account for the experiences of most instructors in IBL courses.

There are two primary sources of data collected for this study: interviews with instructors and documents they submitted (e.g., course syllabi, homework assignments, exams). The semi-structured interview protocol used in the interviews covered many areas. The analysis reported here focuses on five questions: "What are your goals for the students?" "What skills do you expect your students to develop?" "How do you know students are learning?" "How do you measure that students are learning?" and "What type of feedback do you give students?" The first two questions are used to identify the objectives of the instructor. The third and fourth questions address how progress towards these objectives is assessed. The last question examines the response that is given to the student after the assessment is conducted.

With the documents, we explore the quality of the tasks assigned (White, Blum, & Mesa, 2013), as well as the alignment between the instructors' goals and the methods of assessing students learning. The transcripts of the interviews have been analyzed using an open-coding process (Corbin & Strauss, 2008) to identify themes common to instructors' responses. Frequencies of the occurrence each code were calculated, both with respect to the level of instructors' familiarity with IBL methods and to the type of course. This allowed us to identify commonalities in the instructors' objectives, the methods used to assess students' progress toward these objectives, and the feedback given to the students after these assessments were completed.

Findings

We report findings from the interview analysis. Themes that we identified are presented in Table 1. These themes and sub-themes were identified when coding the instructors' responses to the five questions listed above.

Table 1. Themes and Examples about Assessing Students

Category	Examples
Duration	Timed, untimed, revisions, schedule
Formality	Formal, semi-formal, informal
Location	In-class, take-home, on-going
Participant	Individual, group, transition
Process	Designing, conducting, grading, responding
Method	Homework, exams, presentations, discussions
Reason	Formation, summative, instructional
Objectives	Communication, content mastery, disposition

When asked about their objectives for the course, instructors gave formative and summative targets such as wanting students to be able to communicate effectively, wanting students to take responsibility for their learning, and wanting students to cover the necessary content. Though covering the necessary material for a particular class was an underlying goal for many instructors (39%), most instructors (65%) also named process skills (e.g., problem solving, reasoning, multiple representations) as more fundamental goals for their students. One instructor recognized that content knowledge was a goal, but stated that it was "usually secondary to those… process goals" (Albert Austin[1], line 232). Many instructors (78%) cited participation in discussions and clear presentations as evidence they used to know that students are learning. Similarly, responses to the question about skills that students might develop focused on multiple methods of problem solving and building a positive attitude about mathematics. Speaking about a course for future teachers, one instructor said, "these people are going to be teachers and they need to understand other ways of approaching things" (Chelsea Biff, lines 248-250). Taken as a whole, instructors reported expecting their students to learn to communicate well, to solve problems, to adopt a positive attitude about mathematics, and take responsibility for their progress.

[1] Pseudonyms.

When asked about how they *know* their students are learning, the majority of instructors (83%) listed methods used in formative assessment (e.g., monitoring communications, presentations, questions, and discussions). When asked about how they *measure* that students are learning, instructors cited summative methods (e.g., homework, quizzes, tests, exams), with 87% reporting using some of these methods. Though many instructors (43%) spoke about how they distribute scores, points, and grades, one instructor spoke about feeling unconfident, saying he was "not quite sure [he felt] entirely ready to judge their learning process by just seeing what they do on the board" (Adam Bloom, lines 563-564).

When asked about what type of feedback they give to students after these assessments, instructors spoke about the "extensive comments" and the different forms the feedback might take. The time required to give thorough responses was frequently mentioned (. One instructor said she does not "always feel like [they] have time to give as much feedback as I might otherwise want to" (Alicia Biff, lines 758-59). Another instructor also spoke about the time that goes into this teaching method, saying "every assessment took tons and tons of hours…to grade…that was a big part about this method of teaching…you have to give back feedback with another question" (Addison Austin, lines 784-787).

Instructors in the sample identified improving communication skills as a goal for their students. Despite this goal, the techniques they reported using to assess their students are much more summative and resemble the final examinations that are conducted in lecture-based courses.

We have identified differences with respect to the instructors' familiarity with IBL teaching methods and with respect to the type of course they are teaching. Instructors with more familiarity seem less likely to return to "traditional" assessment methods such as thoroughly grading homework assignments and conducting final exams. Instructors in courses for future teachers cited the need to be comfortable with many different solution methods as students in their classrooms may not all use the same method. In addition, the majority of these instructors (75%) put a strong emphasis on students developing a positive disposition towards mathematics.

Discussion

The goals and skills that instructors set for students to reach and develop are similar to the attitudinal benefits that Laursen and Hassi (2010) have documented for IBL courses. Many instructors, however, described the measurement of learning as "like a traditional class" (Allison Bloom, line 623). It aligns more closely with the summative assessment methods used in lecture-based courses and methods students may have experienced in other courses. It is interesting that these instructors reported having non-traditional goals for their students but reported using more traditional assessment methods. This is consistent with Wiliam's (2007) work that found that, when asked about knowing that students are learning, teachers often list informal methods such as classroom questions, group work, and discussions, but, when they are asked about measuring this learning, they list more formal methods, such as tests, homework assignments, and graded projects. This lack of alignment between teaching and assessing may confuse students who have adapted to a style of teaching quite different than previously attended lecture-based courses.

Questions for the audience

1. The NCTM standards provide an accessible framework of assessment in K-12 mathematics. Are there standards more specific for post-secondary students that address the same issues?

2. Do the themes presented in Table 1 make sense? Are they comprehensive? Redundant?

References

Coppin, C. A., Mahavier, W. T., May, E. L., & Parker, G. E. (2009). *The Moore Method: A pathway to learner-centered instruction*. Washington, DC: The Mathematical Association of America.

Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 3rd ed. Los Angeles, CA: SAGE Publications, Inc.

Davis, G., & McGowan, M. (2007). Formative feedback and the mindful teaching of mathematics. *Australian Senior Mathematics Journal, 21*(1), 19-29.

Iannone, P., & Simpson, A. (2011). The summative assessment diet: how we assess in mathematics degrees. *Teaching Mathematics and Its Applications, 30,* 186-196.

Kogan, M., & Laursen, S. (2012). Role of Faculty Professional Development in Improving Undergraduate Mathematics Education: The Case of IBL Workshops. Proceedings of the SIGMAA RUME: *Conference on Research in Undergraduate Mathematics Education*. Portland, OR. http://sigmaa.maa.org /rume/crume2012

Laursen, S., & Hassi, M.-L. (2010). *Benefits of inquiry based learning for undergraduate college mathematics students.* Paper presented at the Annual meeting of the American Educational Research Association, Denver, CO.

Maaß, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching—a synthesis. *ZDM—The International Journal on Mathematics Education*, *45*(6), 779-795.

Madison, B. L. (2006). Tensions and tethers: Assessing learning in undergraduate mathematics. In Steen, L. A. (Ed.) *Supporting assessment in undergraduate mathematics.* (pp. 3-10). Washington, D. C.: Mathematical Association of America.

Noss, R., Goldstein, H., & Hoyles, C. (1989). Graded assessment and learning hierarchies in mathematics. *British Educational Research Journal, 15* (2), 109-120.

Stull, J., Varnum, S., Ducette, J., Schiller, J., & Bernacki, M. (2011). The many faces of formative assessment. *International Journal of Teaching and Learning in Higher Education, 23*(1), 30-39.

Torrance, H. (2007). Assessment *as* learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education, 14* (3), 281-294.

Webb, N. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. A. Grouws (Ed.) *Handbook of research on mathematics teaching and learning* (pp. 661 – 683). Reston, VA: National Council of Teachers of Mathematics.

White, N. J., Blum, C., & Mesa, V. (2013). *Task analysis across different contexts in calculus I*. Paper presented at the Psychology of Mathematics Education-North American Chapter, Chicago, IL.

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F. K. Lester, Jr. (Ed.) *Second handbook of research on mathematics teaching and learning* (pp. 1053-1098). Charlotte, NC: Information Age Printing.

Yoshinobu, S., & Jones, M. (2012). The coverage issue. *PRIMUS, 22,* 303-316.

COMPARING CALCULUS STUDENTS' REPRESENTATION USE ACROSS INTERVIEW AND IN CLASS GROUP-WORK SETTINGS

Dov Zazkis
Rutgers University

The distinction between analytic (notation-based) and visual (diagram-based) representations within students' mathematical problem-solving has been part of the cognitive phycology and mathematics education literature for more than 40 years. However, in spite of this long history there are many unanswered questions regarding how and why particular students choose particular representations, and what influences their social surroundings have on their individual representation use. This study coordinates analyses of calculus students' analytic and visual reasoning across both one-on-one interview and group-work settings. This analysis helps clarify differences between individual representation use and representation use in group settings.

Key words: [calculus, commognitive perspective, functions, visual/analytic reasoning]

The distinction between analytic and visual reasoning can be traced at least as far back as the work of Krutetskii (1976). He established a distinction between analytic thinkers, who prefer to reason in verbal and logical ways, visual thinkers, who prefer the use of visual and spatial reasoning, and harmonic thinkers, who regularly employ both types of reasoning. The use of these and similar distinctions has a long history. Compatible distinctions include: graphical thinking/analytic thinking (Vinner, 1989), visualizers/nonvisualizers (Presmeg, 1986, 1992), depictive/descriptive (Schnotz, 2002) and semantic/syntactic (Weber & Alcock, 2004).

The research based on these distinctions between representations can be partitioned broadly into two categories, one focused on individual cognition and the other focused on representations in multi-person contexts. A long line of individual centered research has examined the prevalence of visual vs. analytic reasoning in various groups. This includes foci such as mathematical giftedness (Presmeg, 1986), differences in sex (George, 1999), teachers' explanations (Gray, 1999) and cultural differences (Presmeg & Bergsten, 1995). Several authors have also modeled student thinking as a process facilitated by transitions between analytic and visual reasoning (Duval, 1999; Stylianou, 2001; R. Zazkis, Dubinsky, & Dautermann, 1996; Zazkis, 2013).

The social centered research on representations has a shorter history than cognitive explorations of this phenomenon. However, a number of important results have been established. This includes that the meaning of representations shifts over time, that children are able to generate, compare, refine and choose amongst competing representations and that novices may under certain circumstances display more sophisticated representation use than experts (Cobb, 2002; diSessa, & Sherin, 2000; Meira, 1998; Roth 2009; Roth & McGinn, 1998).

Both the social and cognitive research on representation use have remained largely disjoint. This is particularly evident in reviews of representation use which often focus on one body of work while ignoring the other (e.g., Bishop, 1989; Hana & Sidoli, 2007; Presmeg, 2006; Roth & McGinn, 1998).

This study aims to bridge these two bodies of work by attempting to tie this study to both of these bodies of work and use both to inform the phenomena observed here. Such bridging is not new (e.g. Cobb & Yackel, 1996), however, it has not been a major theme in literature on representation use.

The Study.

This study followed a group of three students enrolled in a technologically enriched calculus class in a large southwestern university in the United States. These students were given the gender preserving pseudonyms, Ann, Brad and Carson. These students were video recorded during both in class group-work and through a series of three one-on-one problem-solving interviews. The entirety of the interviews and selected days from the in class group-work were then transcribed and coded using the analytic/visual distinction. Physical reasoning was also addressed within the coding (Zazkis, 2013). However, this is not a theme here. Episodes dealing with compatible mathematical subject matter from both settings were then compared. This as well as comparisons of overall representation use patterns facilitate the analysis discussed here.

Results

Comparing the three students' uses of visual, analytic and physical reasoning across interview and group work yielded some interesting parallels. This includes similarities in which modes of reasoning emerged, how often they emerged and problem dependent representation behavior. Cognitive models of representation usage predict such parallels in student reasoning behavior across settings. However, how individual representation use is affected by ones social surroundings is not well understood. Those studying this phenomena from cognitive perspectives have theorized about social influences on individual representation use, but spent little time examining this phenomena. Those examining the phenomena from social perspectives have largely avoided examining individual representation use within multi-person settings and have instead focused on the role representations play in mediating conversations.

In the following I present two tasks. The first was given during in class group work and the other during one-on-one interviews. The comparison of transcripts from the two tasks helps provide incites into how individual representation use is affected by social interactions.

Periodic Function Tasks

In this short paper I will compare student work on two tasks, both of which deal with sine and cosine functions. The first task, known as the tangent intuition task, was used to introduce the concept of graphing derivative. Students worked on this task in small groups in class. Students were given a series of functions represented with both analytic and visual representations and asked to produce the derivative of these functions in both representations. One of these functions was sine. The second task was given during one-on-one interviews with students. Students were given three periodic functions along an unlabeled axis (Figure 1). They were asked to determine the derivative relationship among these functions. Transcripts from both of these will be discussed and compared in the next section.

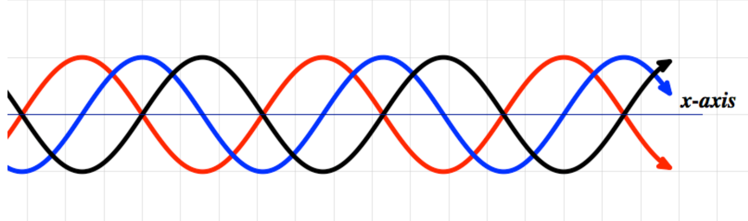


Figure 1: Periodic function task diagram

Comparing Two Episodes

The following episode occurred during group-work on the periodic function task (Figure 1):

Excerpt 1a:

[00:35:40] Carson: Well I would first label the black one as my f of x and then I can see that the blue one is f prime and then the red one is f double prime.

[00:36:14]Researcher: So could you tell me a little bit about how you just did that.

[00:36:18]Carson: So if you're looking at the black one you can see that this is your, right there and there are your critical points and then later this is your point of inflection and then for f prime you can see that there and there and there turn out to be your roots and then the max I mean the critical point turns out to be there which is the critical point, which is the point of inflection from your f of x. And then so if you want to look at the red, the red from prime to double prime you see that the blue one here your point of here's your critical point and then here is your other critical point and this is your point of inflection. So then from blue to red you have your zeros which is here which is you know from your stationary point. And then later here is your second stationary point or critical point. And this is your point of inflection, the blue one so this is your next, then that turns out to be your critical point. Of the second derivative. Yeah.

Carson begins by introducing a solution and expressing it in terms of analytic notation to identify derivative relationships. This solution provided little detail other than Carson's final outcome. The interviewer's prompt for further explanation, "So could you tell me a little bit about how you just did that," does not add new content to the conversation. The interviewer's prompt asks Carson to elaborate on his reasoning. This elaboration remains in a visual mode as opposed to the analytic notation he used to label the functions in his initial stating of his solution. Carson reasons that the maxima and minima of the black function correspond to zeros of the blue function and that the points where the function changes concavity correspond to the maxima and minima of the red function. So although Carson identifies derivative relationships using analytic notation when reasoning about these relationships he uses language that refers to visual representations.

Also notice, that he has not identified the function as a known graphical object at this point, and has only worked with the functions as generic sketches of periodic functions. In the continuation of the transcript this behavior shifts after an interviewer question:

Excerpt 1b:

[00:37:54]Researcher: Alright, are there any other ways you could have approached the task and come to the same conclusion?

[00:38:06]Carson: Is there any other way that I could have approached… I mean well this looks like a sine cosine graph so I would write myself the equation f of x equals I'm going to say sine. And there is also a cosine in here so I guess you can write f of x equals cosine and then and then there's another one which is which looks like x is negative sine [writes f(x)=-sin(x)].

[00:38:50]Researcher: Wait, are all these F's the same F?

[00:38:55]Carson: No, but this line and this line and this line, This [points a f(x)=-sin(x)] would be your black one, and this [writes f'(x)=cos(x)] would be your blue one and this [f"(x)=sin(x)] would be your red one. And so there relationship with each other is, the derivative of sine is cosine and the derivative of cosine is negative sine. And so since this red one right here looks the most like sine I would give this label f of x to sine. And then later…

Notice that in the above episode Carson only transitions to analytic reasoning after being prompted to do so by the interviewer when he says "are there any other ways you could have approached the task?" Such uses of analytic reasoning did not occur within Carson's work without similar prompts from interlocutors. This points to an attribute of Carson's personal representation use that is tied to interactions with others.

Another compatible episode is explored below. The episode occurred during in-class group work on the day that graphing derivative was first introduced in class.

Excerpt 2:

*[00:15:31]Carson: slope is zero here. Zero here and zero here [moving along sin and pointing at the maxima and minima] zero here. This is decreasing, increasing, decreasing, [running pen along function saying increasing when slope is positive and decreasing when negative] No wait. Decreasing increasing…. Yea so that's why I thought [sketches a drawing that looks like 1/6*cos(x)].*
[00:16:34]Brad: Isn't this just negative cosine [pointing at Carson's graph]
[00:16:44]Carson: No, or is that [inaudible]. Yeah it actually is.
[00:16:51]Brad: The derivative of cosine, which is sine.
[00:16:56]Researcher: But they gave you sine.
[00:17:01]Brad: The derivative of sine is negative cosine [writes sin(x)->-cos(x)].
[00:17:10]Carson: Yeah I think it's negative cosine... Well it's positive [pointing to Brad's notation] .
[00:17:26]Researcher: Negative cosine goes through negative one.
[00:17:32]Carson: Cosine is negative sine and … is that right?
[00:17:35]Brad: Is cosine negative sine and sine positive cosine?[writes -cos(x)->sin(x) and changes the – in his previous notation to a +]
[00:17:43]Carson: Yeah, Alright tangent is one over cosine squared [writes tan(x)->1/cos^2(x)]

Notice that here too the interaction shifts to discussion of periodic functions in terms of known analytic referents. Brads question in the above transcript "Isn't this just negative cosine", and the interviewer question in the previous transcript, "are there any other ways you could have approached the task?" point to similar roles. In both cases an interaction that begins with Carson reasoning in the visual mode about slope shifts to an analytic reasoning. So the prompt for a new formulation in the interview and the addition of a new formulation in the group-work both catalyzed similar shifts in representation usage. In both of the above episodes the shift to using analytic reasoning carried through to the subsequent similar task.

Discussion

Heinze, Star and Verschaffel (2009) in their commentary in a recent ZDM special issue on flexible and adaptive use of strategies and representations in mathematics education noted that there are many unanswered questions in literature on representation use. This includes which individual and non-individual centered factors influence representation use, the types of instruction that yield flexible an adaptive representation use and the creation of models that predict representation use. Although, it is clear that social surroundings influence individual representation use the mechanisms that facilitate this influence are poorly understood.

In this paper I have described one such mechanism, prompts for alternate representations. This is an important element to include in future models of representation use. In the two episodes discussed above such prompts caused shifts in individual representation use. These two episodes occurred in different settings but manifested in similar ways. Carson's prolonged use of analytic reasoning did not occur without such prompts. These episodes reveal that Carson is quite adept at analytic reasoning, but does not often use such reasoning without the prompting of others. So preference and use of one representation in lieu of another is not strictly tied to a student's ability to reason with that representation.

Additionally, other study participants did not demonstrate similar shifts in representation use in response to compatible prompts for alternate methods. So these prompts do not influence all students in the same way. Instead this data helps illustrate that individual representation use and how it changes in response to interlocutors differs among students. However, there is, at least with in the data here, a fair amount of consistency interns of how students approach tasks and how these approaches shift in response to particular prompts.

References

Cobb, P. (2002). Reasoning with tools and inscriptions. *Journal of the Learning Sciences*, *11*(2-3), 187-215.

Cobb, P. & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, 31, 175–190.

Disessa, A. A., & Sherin, B. L. (2000). Meta-representation: An introduction. *The Journal of Mathematical Behavior, 19(4), 385-398.*

Duval, R. (1999). Representation, Vision and Visualization: Cognitive Functions in Mathematical Thinking. Basic Issues for Learning.

George, E. A. (1999). Male and female calculus students' use of visual representations. In O. Zaslavsky, (Ed.), *Proceedings of the 23rd PME International Conference, 3,* 17-24.

Gray, E. (1999). Spatial strategies and visualisation. In *Proceedings of XXIII International Conference for the Psychology of Mathematics Education, Haifa, Israel*.

Hanna, G. & Sidoli, N. (2007). Visualisation and proof: a brief survey of philosophical perspectives. *ZDM,* 39, 73–78

Heinze, A., Star, J. R., & Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM*, *41*(5), 535-540.

Krutetskii, V. A. (1976). *The Psychology of Mathematical Abilities in Schoolchildren*, Chicago: University of Chicago Press.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. Journal for Research in Mathematics Education, 121-142.

Presmeg, N. C. (1986). Visualization and mathematical giftedness. *Educational Studies in Mathematics, 17*, 297-311.

Presmeg, N. C. (2006). Research on visualization in learning and teaching mathematics. *Handbook of research on the psychology of mathematics education*, 205-235.

Presmeg, N. C., & Bergsten, C. (1995). Preference for visual methods: An international study. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19th PME International Conference, 3*, 58-65.

Roth, W. M. (2009). *Mathematical representation at the interface of body and culture*. Iap.

Roth, W. M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of educational research*, *68*(1), 35-59.

Schnotz, W. (2002). Commentary e Towards an integrated view of learning from text and visual displays. *Educational Psychology Review,* 14(1), 101-120.

Stylianou, D. (2001). On the reluctance to visualize in mathematics: Is the picture changing? In M. vanden Heuvel-Panhuizen (Ed.), *Proceedings of the 25th PME International Conference, 4*, 225-232.

Vinner, S. (1989) The avoidance of visual considerations in calculus students. *Focus on Learning Problems in Mathematics,* 11(2), 149-156.

Zazkis, R., Dubinsky, E., & Dautermann, J. (1996). Coordinating Visual and Analytic Strategies: A study of students' understanding of the group D 4. *Journal for research in Mathematics Education*, 435-457.

Zazkis, D. (2013). Calculus Students' Representation Use in Group-Work and Individual Settings. (Unpublished doctoral dissertation). San Diego State University, San Diego.

PROOF SCRIPTS AS A LENS FOR EXPLORING PROOF COMPREHENSION

Rina Zazkis Dov Zazkis
Simon Fraser University Rutgers University

We examine perspective secondary teachers' conceptions of what constitutes comprehension of a given proof and their ideas of how students' comprehension can be evaluated. These are explored using a relatively novel approach, scripted dialogues. The analysis utilizes and refines Mejia-Ramos, Fuller, Weber, Rhoads and Samkoff's (2012) proof comprehension framework. We suggest that this refinement is applicable to other studies on proof comprehension.

Key words: Proof; Proof comprehension; Proof scripts; Pythagorean theorem

PROOF COMPREHENSION

When undergraduate mathematics courses involve proofs, a request to prove a theorem or a related statement is a common assessment method used in examinations. However, this method often focuses on students' rote learning and memorization rather than their grasp of a theorem (Conradie & Frith, 2000). To address the issue Conradie and Frith (2000) introduced a 'comprehension test', a method in which a proof of some result is presented to students and they are asked to answer questions related to particular claims within the presented proof. They exemplified the method using two different proofs and noted possible modifications and extensions of the method. One such modification was related to filling gaps in the presented proofs. It was further suggested that comprehension tests provide "a far more precise evaluation of a student's understanding at all levels" (ibid, p. 231), as well as improved feedback about student learning.

Yang and Lin (2008) claimed that despite the centrality of understanding in mathematics education, reading comprehension is underemphasized in proof instruction. In their study of proof in a geometry setting, they designed a hierarchical four-level model of proof comprehension. The first *surface* level attends to student understanding of particular statements and symbols used in the proof, without focusing on how particular statements relate to each other. The *recognizing the elements* level attends to the logical status of the statements that appear in the proof, either explicitly or implicitly. The next level, *chaining the elements*, focuses on connections among different statements. At the final level four, *encapsulation*, students reflect on the proof as a whole in terms of main ideas and methods and may consider its application in other contexts. However, because their investigation took place in a school geometry setting the focus of these researchers was on the first three levels.

Mejia-Ramos, Fuller, Weber, Rhoads and Samkoff (2012) created an assessment model of proof comprehension. This model extended and refined Yang and Lin's (2008) model and provided detail with regard to how students' comprehension might be evaluated. This model helped delineate the types of items that may appear on comprehension tests, such as those discussed by Conradie and Frith (2000). Due to its focus on undergraduate proof, the Mejia-Ramos et. al. (2012) framework further delineated Yang and Lin's (2008) level four, *encapsulation*. The Mejia-Ramos et. al. framework separates holistic comprehension, which includes main ideas and proof methods, and local comprehension, which includes a logically derived series of steps. The first three of their categories correspond to local comprehension (1. Meaning of terms and statements; 2. Logical status of statements and proof framework; and 3. Justification of claims). The last four correspond to holistic comprehension (4.Summarizing via high-level ideas; 5. Identifying modular structure; 6.Transferring the general ideas or methods to another context; and 7. Illustrating with examples).

While Mejia-Ramos et al. (2012) considered their adaptation and reinterpretation of levels relevant to proof comprehension in undergraduate mathematics, we find elements of their model relevant and useful in discussing comprehension of a relatively uncomplicated proof from school mathematics. We elaborate upon these elements in a next section.

PROOF SCRIPTS

The proof script method, which involves presenting a proof in a form of a scripted dialogue, was inspired by Lakatos's evocative "Proofs and Refutations". The roots of this method can be further traced to a Socratic dialogue, where communication among characters helps in uncovering and resolving flaws or inconsistencies in one's thinking.

Exploring conceptions of instructional interaction through scripts is featured in the studies of Zazkis, Sinclair and Liljedahl (2013), who introduced the construct of a 'lesson play', which is, an imagined interaction among a teacher and her students, presented in a form of a script. Zazkis et al. (2013) analyzed lesson plays composed by prospective teachers on a variety of topics related to elementary school mathematics. They argued that asking prospective teachers to think about their future teaching in terms of fictional interactions draws their attention to how their students' mathematical thinking can be developed. They described the affordances of this approach both in teacher education and in research. In teacher education it provided a valuable tool for engaging prospective teachers in considering particular students mistakes or difficulties. In research it provided a lens for exploring how prospective teachers envision addressing students' difficulties, both mathematically and pedagogically. In particular, the prospective teachers' personal understanding and conceptions of the mathematics involved became apparent in their attempts to guide students' solutions.

Similarly, the mathematical understandings of script-writers can become visible when they attend to particular proofs, rather than to instructional interaction in general. We use the term 'proof scripts' to refer to scripted dialogues that elaborate on mathematical proofs. Several studies engaged participants in producing such proof scripts. In Gholamazad (2006, 2007) prospective elementary school teachers interpreted basic proofs in elementary number theory, such as, if a divides b, and b divides c, then a divides c. Prospective secondary school teachers in Koichu and Zazkis (2013) elaborated upon Fermat's Little Theorem. The study of Zazkis (2013) used the script-writing method in considering a proof that derivative of an even function is odd. The method proved fruitful in identifying participants' ideas with respect to the key elements in the given proofs and potential difficulties in understanding the proofs.

This study uses proof scripts as a data collection method. However (unlike prior studies where the choice of the characters in proof scripts was left to script writers) it introduces a constraint: the characters are a teacher and a student, and the purpose of the dialogue is for the teacher-character to assess the student's comprehension, predict possible pitfalls and help the student overcome them. The resulting proof scripts give insights into participants' own conceptions of what ideas are central to the given proof as well as to their conceptions of proof assessment.

THE STUDY

Our study centers on comprehension of a particular proof of the Pythagorean theorem, presented in Figure 1. Participants were 24 prospective secondary school mathematics teachers in their final term of a teacher education program. The participants held degrees in mathematics or science and at the time of data collection were enrolled in a problem solving course, one of the goals of which was to deepen their knowledge of school mathematics. They were asked to respond in writing to the Task below.

Consider the following proof of the Pythagorean theorem

Draw a square ABCD in which the length of the side is a+b.

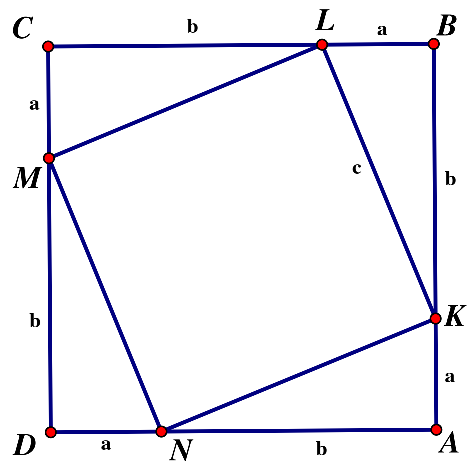
Connect points KLMN.

The area of ABCD is $(a+b)^2$

However, this area can also be calculated as composed of the square KLMN and 4 triangles, that is,

$4 \times \frac{1}{2}(ab) + c^2 = (a+b)^2$
$2ab + c^2 = a^2 + 2ab + b^2 \implies$
$c^2 = a^2 + b^2$

QED



Imagine that you are working with a high school student and testing his/her understanding of different aspects of this proof.

What would you ask? What would s/he answer if her understanding is incomplete? How would you guide this student towards enhanced understanding? Identify several issues in this proof that may not be completely understood by a student and consider how you could address such difficulties. In your submission:

(a) Write a paragraph on what you believe could be a "problematic point" (or several points) in the understanding of the theorem/statement or its proof for a learner.

(b) Write a scripted dialogue between teacher and student that shows how the hypothetical problematic points you highlighted in part (a) could be worked out (THIS IS THE MAIN PART OF THE TASK).

(c) Add a commentary to several lines in the dialogue that you created, explaining your choices of questions and answers.

Figure 1: The Task

In short, the participants were asked to consider the given proof of the Pythagorean theorem and write a script for a conversation between a teacher and a student, in which the teacher-character assesses the student's understanding of the proof. The participants were further advised that the dialogue should expose their views of what could be problematic for a student, that some issues in the proof may not be completely understood, and that the role of the teacher-character was to uncover and address the student-characters' difficulties.

The script composed by participants can be seen as an imagined oral test of proof comprehension. The following **research questions** guided our analysis: (a) What are the participants' ideas regarding what is important to understand in the given proof, (b) What do they perceive as potential difficulties for students, and (c) What are their conceptions of how proof comprehension can be evaluated.

A-PRIORI TASK ANALYSIS

Using the relevant elements from Mejia-Ramos and his Colleagues' framework (MRC) we constructed a proof comprehension test (Conradie & Frith, 2000). That is, we listed what questions would be used if we were to complete the Task assigned to our participants ourselves. Given the page limit, we focus here on MRC type #3, *justification of claims*:

- It is claimed that KLMN is a square, why?
- It is assumed that the 4 triangles have the same area, why?
- How is the equation – $4 \times ½(ab) + c^2 = (a+b)^2$ –derived?
- How is the equation manipulated and simplified?

DATA ANALYSIS

Given the relative mathematical maturity of our participants, we were not interested in their personal responses to the questions in our comprehension test, but in their ways of assessing and ensuring students' comprehension. Our tenet is that by questioning a student – or designing an imaginary dialogue with a student – participants expose their personal attention to the various elements of the proof and the importance of these elements in students' proof comprehension.

As noted in our 'comprehension test', the *justification of claims* aspect of local comprehension (MRC, #3) involves interpretation of the geometric figures in the diagram and of the algebraic formulas. However, algebraic manipulation was central to most of the proof scripts, while the geometry of the situation was either ignored or treated only partially.

Attention to Algebra. Most scripts (18 out of 24) included an explanation of how the binomial $(a+b)^2$ is expanded. In 11 of these the explanation attended to a potential student's difficulty in manipulating $(a+b)^2$ and erroneously equating it to $a^2 + b^2$,which is problematic in the last steps of the proof. These scripts introduced the student's error and then corrected it in various ways. For example, participant #16 (P-16) in his script addresses the student-character's error with an invitation to consider a numerical example. This results in immediate correction, likely in reference to a previously learned rule.

[P-16.1] Teacher:	Hmm, well is $(a + b)^2 = a^2 + b^2$? Why don't you try this with a = 3 and b = 4.
[P-16.2] Student:	Okay. $(3 + 4)^2 = 49$. And $3^2+ 4^2 = 25$. Whoa, that's not right. Oh! I should have foiled!
[P-16.3] Teacher:	Correct.

Of note is that the student's claim "I should have foiled" [P-16-2] is an example of verbification of a commonly used acronym FOIL (first-outside-inside-last). This is a mnemonic device used to help in remembering rules for multiplying binomials.
In P-10 the correct formula suggested by the teacher is confirmed by a numerical example.

[P-10.1] Teacher:	Well let's expand and see if we can simplify the equation. $(a+b)^2$ is equal to $a^2 + 2ab + b^2$.
[P-10.2] Student:	Why would that be, shouldn't it equal $a^2 + b^2$?
[P-10.3] Teacher:	Let's examine that idea by substituting in numbers for a and b. Have a = 2 and b = 3. If $(a+b)^2$ is equal to $a^2 + b^2$ then $(2+3)^2$ would equal $2^2 + 3^2$ or 4 + 9 = 16. Is this correct?
[P-10.4] Student:	Well if we use the BEDMAS rule and do the brackets first 2+3 = 5 and 5 squared is 25, so that isn't right.
[P-10.5] Teacher:	Remember that $(a+b)^2$ means (a+b)(a+b), it is the area of square ABCD (with lengths a+b). Remember to use the foil method when you see this type of expression to ensure that you are multiplying all the variables together. Try substituting 2 and 3 into $a^2 + 2ab + b^2$ and see if you get the correct answer.
[P-10.6] Student:	Okay so $2^2 + 2(2)(3) + 3^2 = 4 + 12 + 9 = 25$. I see now.

In summary, the majority of proof-scripts attended to algebraic manipulations within the proof. Inappropriate expansion of binomials, as an example of overgeneralized linearity, was identified by the script-writers as the main problematic point for student-characters. Correct expansion was achieved by either recalling the formula, considering numerical examples or by geometric demonstration.

Attention to Geometry. Of the 24 proof scripts created by the participants, 14 paid no attention to justifying the geometric shapes and relations in the given proof. One possible explanation for the partial treatment of geometric aspects within participants' proof scripts can be found within the commentary of participant #12 (P-12) that accompanied her script. We agree with the view expressed in the following excerpt:

> *Most students can describe a square as having four equal sides but they often forget about the 90 degree angles. Without this fact the rest of the proof would be difficult if not impossible. However, the most problematic point about this proof is the assumption that KLMN is a square. Many students will skim over this line of the proof and assume it to be true because it says that KLMN is a square.*

In fact, not only students "will skim over this line", as suggested by P-12, but teachers participating in our study did so as well. In our data, only two participants provided complete mathematical justifications for KLMN being a square in their scripts. In eight other cases the square was acknowledged, but the provided justification included attention only to the equal sides. This is exemplified in the excerpt from participant #20 (P-20):

[P-20.1] Teacher:	Let's label it KLMN. Do we know for sure if it is a square?
[P-20.2] Student:	We have to check if each side is the same length.
[P-20.3] Teacher:	Right, how is each length in KLMN related to a and b?
[P-20.4] Student:	The three sides create a triangle.
[P-20.5] Teacher:	Great. And the side we don't know is called?
[P-20.6] Student:	The hypotenuse.
[P-20.7] Teacher:	Good.
[P-20.8] Student:	And since each side of KLMN is the hypotenuse of a triangle with sides a and b, then KLMN is a square.
[P-20.9] Teacher:	Yup, so KLMN is a square. As I said at the start, we want to derive the Pythagorean theorem using the area of ABCD. What is the area of a square or rectangle in general?

Here, "each side is the same length" [P-20.2] is suggested by a student as a sufficient property for a square and is approved by the teacher [P-20.3]. The side-angle-side congruency property is implied; though the angle of 90 degrees is never mentioned, the right angle triangle is implied when referring to a hypotenuse [P-20.8].

Prospective teachers participating in this study undoubtedly knew the definition of a square and could prove the "squareness" of KLMN without any difficulty. (In-class discussion that followed administration of the Task confirmed this claim). However, in the majority of the scripts we found either no attention to the need for proving this property or erroneous and incomplete justifications. This demonstrates the participants' views on what is important *in the proof* and what is essential *for the student* to understand. We wonder whether the geometric issues in the proof would have been treated differently if mentioned as separate claims in the proof, rather than presented as 'obvious'.

We note that in elaboration on assessment type #3 in MRC, *justification of claims*, the authors note that some warrants in a proof may be implicit, under the assumption that it may be obvious to the reader. However, they make this comment with respect to statements of the form "Since A, then B". In such statements a claim (B) has data to support it (A). With respect to the particular proof discussed here, the claim that KLMN is a square is not only implicit, it is 'concealed'. That is, it does not have either data or a warrant to support it. Instead it appears in the same way data would, as a fact that helps support and argument rather than a claim that may need to be justified. We suggest that attention to such concealed elements—claims that appear as data with no further justification—should explicitly feature in a proof assessment framework.

Conclusion

We explored how prospective secondary teachers of mathematics envision assessing proof comprehension of a given proof of the Pythagorean theorem. Proof-scripts composed by prospective teachers serve as a mirror of their mathematical and pedagogical attention. Extensive elaboration on algebraic manipulations that appear in the proof-scripts can be explained as attention to an explicit claim in the proof and also as attention to a familiar student error. Participants demonstrated a variety of strategies in addressing this error in their scripts, from simple reminders to visual and numerical demonstrations. This points to knowledge of a range of pedagogical approaches as well as knowledge of common student errors.

In this report we focused on comprehension type #3 of the MRC framework, *justification of claims*. The scripts revealed that participants had difficulty distinguishing between data used to support claims and claims themselves. While the participants, when prompted, had no problem proving that KLMN was a square considering both its sides and its angles, the majority did not attend to this claim within their proof scripts. They instead treated this claim as if it was data used within an argument and did not realized that it may call for (non-trivial) justification. Several factors may have contributed to such lack of attention: reliance on the visual diagram and undoubted acceptance of 'concealed' claims in the proof. We suspect that had the presented proof included an explicit statement, such as "connecting in sequence points K, L, M and N with line segments results in a square KLMN", the script-writers could have attended to the shape of KLMN with more diligence. Such a statement explicitly paints squareness as a claim rather than data. The same can be said about the congruence of the four triangles. We believe that identifying claims that proof-writers could have taken for granted, and alerting their attention to such claims is an important component of pedagogy that was not activated in this group of prospective teachers. As such, our modification of the MRC framework attends explicitly to implicit claims in a proof.

This study can be seen as contributing to two arenas: one methodological and the other theoretical. First, with respect to methodology, we expanded the use of script-writing, and specifically proof scripts, to teacher-student interactions that aim at assessing and modifying student comprehension of a particular proof. In accord with the claims of Zazkis et al. (2013), that lesson plays provide a lens for analyzing teachers' ways of addressing students' difficulties, we add that scripts composed around particular proofs zoom in on particular difficulties associated with these proofs. Further, the scripts highlight the participants' choices of issues to be addressed in assessing student proof-comprehension. This provides a window into these prospective teachers' images of proof assessment and, in particular, how these assessments can be implemented and what they focus on.

With respect to theory, we suggest that comprehension type 3 of the MRC framework be partitioned into 3A: Justification of explicit claims, and 3B: Justification of implicit (concealed) claims. We believe that such expansion is applicable to a variety of different proofs and is not limited to proofs appearing in undergraduate mathematics courses or to proofs in geometry.

REFERENCES

Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics, 42*, 225–235.

Gholamazad, S. (2007). Pre-service elementary school teachers' experiences with the process of creating proofs. In Woo, J. H., Lew, H. C., Park, K. S. & Seo, D. Y. (Eds.), Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education (Vol. 2, pp. 265-272). Seoul, Korea: PME.

Koichu, B., & Zazkis, R. (2013). Decoding a proof of Fermat's Little Theorem via script writing. *The Journal of Mathematical Behavior*, *32*(3), 364-376.

Mejia-Ramos, J.P, Fuller, E. , Weber, K., Rhoads, K. & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics, 79*, 3-18.

Yang, K.-L., & Lin, F.-L. (2008). A model of reading comprehension of geometry proof. *Educational Studiesin Mathematics, 67*, 59–76.

Zazkis, D. (2013). Odd dialogues on odd and even functions. *Proceedings of the Conference for Research in Undergraduate Mathematics Education*. Denver, CO.

Zazkis, R., Sinclair, N., & Liljedahl, P. (2013). *Lesson Play in mathematics education: A tool for research and professional development*. New York, NY: Springer.