PROCEEDINGS OF THE 15$^{\text{TH}}$ ANNUAL CONFERENCE ON RESEARCH IN UNDERGRADUATE MATHEMATICS EDUCATION

EDITORS
STACY BROWN
SEAN LARSEN
KAREN MARRONGELLE
MICHAEL OEHRTMAN

PORTLAND, OREGON
FEBRUARY 23 – FEBRUARY 25, 2012

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

CITATION: In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education*, 2012, Portland, Oregon.

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 fifteenth annual Conference on Research in Undergraduate Mathematics Education in Portland, Oregon from February 23 - 25, 2012.

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. Alan Schoenfeld, Dr. Chris Rasmussen, Dr. Lara Alcock, and Dr. Cynthia Atman, a special session by Dr. Jacqueline Dewar, and the presentation of over 100 contributed, preliminary, and theoretical research reports. In addition to these activities, faculty, students and artists contributed to an inaugural display on Art and Undergraduate Mathematics Education.

The Proceedings of the 15th 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.

Volume 1, *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. Volume 1 begins 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 insights into existing research programs.

Volume 2, *RUME Conference Reports*, includes 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,
Stacy Brown, RUME Conference Chairperson
Karen Marrongelle, RUME Conference Local Organizer
Sean Larsen, RUME Program Chair
Michael Oehrtman, RUME Coordinator

PROGRAM COMMITTEE

Program Committee Chair	
Stacy Brown	Pitzer College
Local Organizer	
Karen Marrongelle	Portland State University
Karen Allen-Keene	North Carolina State University, Raleigh
Jason Belnap	University of Wisconsin Oshkosh
Nicole Engelke	California State University, Fullerton
Tim Fukawa-Connelly	The University of New Hampshire
Evan Fuller	Montclair State University
Hope Gerson	Brigham Young University
Sean Larsen	Portland State University
Michael Oehrtman	University of Northern Colorado
Vicki Sealey	West Virginia University
Craig Swinyard	University of Portland
Allison Toney	University of North Carolina–Wilmington
Joseph Wagner	Xavier University
Aaron Weinberg	Ithaca College

WITH MUCH APPRECIATION WE THANK THE FOLLOWING REVIEWERS

Belnap, Jason (PC)
Brown, Stacy (PC)
Engelke, Nicole (PC)
Fukawa-Connelly, Tim (PC)
Gerson, Hope (PC)
Keene, Karen (PC)
Oehrtman, Michael (PC)
Sealey, Vicki (PC)
Swinyard, Craig (PC)
Toney, Allison (PC)
Wagner, Joe (PC)
Weinberg, Aaron (PC)
Bagley, Spencer
Blaisdell, Rob
Bond, William
Cadwallarolsker, Todd
Capaldi, Mindy
Champney, Danielle
Chesler, Joshua
Code, Warren
Cook, John Paul
Cook, Samuel
Dewar, Jackie
Ellis, Jessica
Ferguson, Leann
Grundmeier, Todd
Halani, Aviva
Hall, William
Jackon, Billy
Johnson, Estrella
Kim, Hyung
Lockwood, Elise
Mantini, Lisa
Mayer, John
Mcallister, Michael
Meija Ramos, Juan Pablo
Mills, Melissa
Moore, Kevin
Park, Jungeun
Rasmussen, Chris
Samkoff, Aron
Savic, Milos
Selinski, Natalie
Sheperd, Mary
Soto-Johnson, Hortensia
Szydlik, Jennifer
Talley, Jana
Wawro, Megan
Weber, Keith
Whitacre, Ian
Zandieh, Michelle

CONFERENCE PARTICIPANT LIST

Adiredja, Aditya
Bagley, Spencer
Baker, Gregory
Beisiegel, Mary
Belnap, Jason
Blaisdell, Rob
Boester, Tim
Bond, William
Brandt, Jim
Brown, Stacy
Bubp, Kelly
Byerley, Cameron
Byrne, Martha
CadwalladerOlsker, Todd
Cakebread, Amy
Capaldi, Mindy
Carlson, Marilyn
Chamberlin, Michelle
Champney, Daniell
Chang, Kuo-Liang
Chernoff, Egan
Chesler, Joshua
Christensen, Warren
Cilli-Turner, Emily
Code, Warren
Conway, Katherine
Cook, John Paul
Cook, Samuel
Cory, Beth
Czeranko, Katherine
Czocher, Jennifer
Dawkins, Paul
Deshler, Jessica
Dewar, Jacqueline
Dibbs, Rebecca
Dray, Tevian
Duranczyk, Irene
Ellis, Jessica
Engelke, Nicole
Ernest, Jessica Brooke
Eubank, Alyssa
Evans, Addie
Ferguson, Leann
Fisher, Brian
Fukawa-Connelly, Timothy
Galle, Gillian
Garrity, Shawn
Giraldo, Jose
Glassmeyer, David
Gleason, Jim
Gol Tabaghi, Shiva
Goldlist, Amy
Grundmeier, Todd
Hachey, Alyse C
Halani, Aviva
Hall, William
Hamdan, May
Hamlin, Alyssa
Harris, Gary
Hatfield, Neil
Hauk, Shandy
Hodge, Angela
James, Kendrice
Jetter, Madeleine
Karakok, Gulden
Katz, Brian
Keene, Karen
Kelton, Molly
Kim, Hyung
Knapp, Jessica
Knight, Brian
Kogan, Marina
Kohler, David
Kustusch, Mary Bridget
Larson, Christine
Laursen, Sandra
Levin, Mariana
Lew, Kristen
Lindaman, Brian
Lockwood, Elise
Lougheed, Tom
Louie, Tammy
Lucas, Timothy
Mamolo, Ami
Mantini, Lisa
Marrongelle, Karen
Martin, Jason
Mayer, John
Mcallister, Michael
Mejia Ramos, Juan Pablo
Mesa, Vilma
Meyers, Beverly
Mickelson, Jesse
Mills, Melissa

Molitoris Miller, Susanna
Moore, Kevin
Mullins, Bernadette
Nemirovsky, Ricardo
Oehrtman, Michael
Park, Jungeun
Pazicni, Samuel
Peters, Elisabeth
Peterson, Scott
Peterson, Valerie
Rand, Betseygail
Raychaudhuri, Debasree
Retsek, Dylan
Rhodehamel, Bohdan
Riley, Douglas
Rimmasch, Gretchen
Robert, Ely
Ross, Rebecca
Salgado, Hilda
Salomone, Stephanie
Samkoff, Aron
Samuels, Jason
Savic, Milos
Sayre, Eleanor
Schmitz, Rebecca
Sealey, Vicki
Seaman, Carol
Selden, Annie
Selden, John
Selinski, Natalie
Shepherd, Mary
Shipulina, Olga
Smith, John
Smith, Michael
Soto, Osvaldo
Soto-Johnson, Hortensia
Stewart, Sepideh
Strickland, Sharon
Subedi, Krishna
Swinyard, Craig
Szydlik, Jennifer
Talley, Jana
Tallman, Michael
Tang, Gail
Theakston, Alexandria
Thomas, Matthew
Titova, Anna
Todd, Rachael
Toney, Allison
Toth, Krista
Trouba, Jerome
Underwood, Kathryn
Vollet, Carly
Wagner, Joseph
Wangberg, Aaron
Watkins, Laura
Wawro, Megan
Weber, Eric
Weber, Keith
Wei, Wei
Weisbart-Moreno, Jennifer
Wheeler, Greg
Whitacre, Ian
White, Nina
Whittemore, Timothy
Wildfeuer, Sandra
Wladis, Claire
Yurasovskaya, Ekaterina
Zandieh, Michelle
Zazkis, Dov
Zazkis, Rina

VOLUME 1

CONFERENCE PAPERS

TABLE OF CONTENTS

Volume 1: Conference Papers

A MODEL OF STUDENTS' COMBINATORIAL THINKING: THE ROLE OF SETS OF OUTCOMES

Elise Lockwood
University of Wisconsin – Madison

Combinatorial topics are prevalent in undergraduate curricula, and research indicates that students face difficulties when solving counting problems. The literature has not sufficiently addressed students' ways of thinking about combinatorial concepts at a level that enables researchers to understand how students conceptualize counting problems. In this paper, a model of students' combinatorial thinking is presented that emphasizes relationships between formulas/expressions, counting processes, and sets of outcomes; additionally, the model is used to frame several examples of students' reasoning about counting problems. The model serves as a conceptual analysis of students' thinking and activity related to counting, providing language to describe and explain aspects of students' counting activity. In this way, the model has practical implications, both for researchers (providing a lens through which to examine data on combinatorics education) and for teachers (providing an aid to instructional design based on student thinking).

Key words: Combinatorics, Counting, Grounded Theory, Model, Discrete Mathematics

Introduction and Motivation

Combinatorics has received increased attention in K-12 and undergraduate curricula (e.g., Batanero, Navarro-Pelayo, & Godino, 1997; English, 1991; NCTM, 2000), both due to its rich potential as a problem solving context, as well as for its applications in probability and computer science. One aspect of combinatorics, counting, is among our earliest intellectual processes. As students advance mathematically, however, they tend to experience considerable difficulty with counting problems (e.g., Batanero, et al., 1997; Martin, 2001). In spite of efforts to improve the implementation of combinatorial topics in the classroom (e.g., Kenney & Hirsch, 1991; NCTM, 2000), students continue to struggle with understanding such concepts. Counting problems can be easy to state, but they can be surprisingly difficult to solve. As we see in reviewing the literature below, more research is needed that explores students' ways of thinking about solving counting problems. In this paper, I share a preliminary model of students' combinatorial thinking that emerged during task-based interviews with post-secondary students.

Literature Review and Guiding Perspectives

Beginning with Piaget's work with young children (Piaget & Inhelder, 1975), which was extended by English (e.g., 1991), mathematics education researchers have since studied a handful of issues related to combinatorics education. Specifically, researchers have identified and suggested variables that may contribute to student difficulties (e.g., Batanero, et al., 1997, Hadar & Hadass, 1987) and have examined students' verification strategies in the domain of combinatorics in particular (Eizenberg & Zaslavsky, 2004). Additionally, another group of researchers have emphasized students' powerful combinatorial potential and have demonstrated how students have successfully made connections across combinatorial ideas (e.g., Maher & Speiser, 1997; Maher, Powell, & Uptegrove, 2011). While this research has been informative, to this point the literature on combinatorics education has not addressed such ways of thinking at a level that enables researchers and educators to understand how students conceptualize counting problems. A fundamental aspect of helping students overcome the difficulty of solving combinatorial problems is to understand students' conceptualizations of such activity. In order to help students succeed combinatorially, researchers need a deeper understanding of students' thinking about the mathematical activity of solving counting problems.

To this end, I conducted research on post-secondary students and attempted to learn more about their ways of thinking about counting problems, and I considered the notion of a conceptual analysis (Von Glasersfeld, 1995) to do so. Thompson's (2008) points out that conceptual analyses can be used in part "to generate models of knowing that help us think about how others might know particular ideas" (p. 57). In this paper, I present such a model of students' combinatorial thinking. By students' combinatorial thinking, I mean my interpretation of their thinking based on their observable words and actions; while I cannot *know* for certain what a student is thinking, I can make inferences based on what they say and do. The word "model" in this context refers to a particular system for identifying, describing, and explaining phenomena related to a particular mathematical topic – in this case combinatorial thinking. I draw upon Lesh and Doerr's (2000) view of models, as they note that, "Not just any old system functions as a model" (p. 362). They go on to say that, "To be a model, a system must be used to describe some other system, or to think about it, or to make sense of it, or to explain it, or to make predictions about it" (p. 362).

The model presented in this paper represents a conceptual analysis of students' thinking and activity related to combinatorial enumeration (counting). It is both empirically and theoretically devised and has been refined and elaborated through analysis of student data. The model is an attempt to make sense of students' combinatorial thinking and to be explanatory, and not just descriptive, in its discussion of significant phenomena. The model serves as a contribution to the field by giving an initial attempt at describing and explaining some ways of thinking that might be beneficial or deleterious for students as they count.

In the remainder of the paper I present the model. I begin by elaborating the specifics of the model, drawing upon illustrative mathematical examples to discuss the components of the model and to examine relevant relationships. I then further explore the model (and make a case for its applicability) by framing some examples from the data in terms of the model. I conclude with a further discussion of the rationale for such a model and describe future research.

Data Collection and Analysis

In order to present the model and to contextualize subsequent discussion of the model, I briefly describe the study from which it emerged. I interviewed twenty-two post-secondary students in 60-90 minute individual, videotaped interviews as they solved five combinatorial tasks. In these interviews, students were encouraged to explain their reasoning as they first solved all five problems on their own, giving their best initial attempt at an answer. Then, I revisited the problems with the students and gave them alternative expressions that I asked them to evaluate. The purpose was to have students make sense of the alternative expression and to make some judgment as to the correctness of their original answer as compared to the alternative expression. During the interviews, my intent was not to instruct the students or to measure learning, but rather to examine their activity and thinking.

Because not much is known about students' ways of thinking about counting, the methodological framework of grounded theory (Strauss & Corbin, 1998) was adopted for the study. At the core of grounded theory is the premise that researchers may study phenomena about which no previously existing theory exists. According to this perspective, raw data is carefully analyzed, relevant phenomena and themes from the data are identified and organized, and theory emerges as the end product of such work. My analysis consisted of transcribing the interviews, searching the transcripts and videos for episodes that highlighted particular phenomena, labeling and structuring the phenomena, and ultimately developing theory out of the analysis process.

Results and Findings

In this section, I introduce the components of the model and describe how they relate to one another. The purpose of this model is to shed light on relevant elements of students' counting and to provide language by which to describe and explain aspects of such counting activity, with the end goal of ultimately highlighting ways in which students might think about combinatorial ideas.

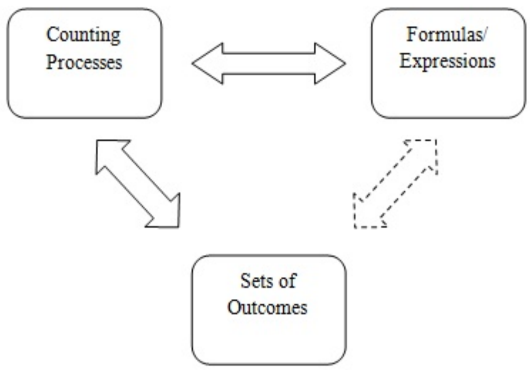


Figure 1: A Model of Students' Combinatorial Thinking

Components of the Model

I begin by explaining each of the components of the model: Formulas/Expressions, Counting Processes, and Sets of Outcomes. *Formulas/Expressions* refer to mathematical expressions that yield a numerical value. The formula could have some inherent combinatorial meaning (such as a binomial coefficient $\binom{8}{3}$), or it could be some combination of numerical operations (such as a sum of products $9\cdot 13+3\cdot 12$). It may be the case that two expressions are mathematically equivalent (in the sense that one expression may be able to be simplified into the other), but they differ in form (that is, the expressions themselves appear different on the page). For this discussion, I consider two expressions to be different if they differ in form. *Counting Processes* refer to the enumeration process (or series of processes) in which a counter engages (either mentally or physically) as he or she solves a counting problem. The implementation of a case breakdown or successive applications of the multiplication principle[1] are examples of counting processes that a counter might enact. *Sets of Outcomes* refer to the objects being counted – those sets of elements that a counter can imagine being generated or enumerated by a counting process.

[1] The multiplication principle states that the number of options for a pair of independent choices is the product of the number of options for each choice (see Martin, 2001 or Tucker, 2002 for more information).

This may be the set whose cardinality represents the answer to that counting problem, but sets of outcomes could also refer to any set that can be associated with a counting process (even if that set is not the answer to the counting problem at hand). For example, in a counting problem asking for the number of 10-letter sequences that contain exactly two consecutive A's, the desirable set of outcomes is all such 10-letter sequences that satisfy the constraint. That set could be considered in light of another set – the set of all possible 10-letter sequences.

Key Relationships between Components of the Model

When working on a given counting problem, a student may draw upon one or more of the above components and may explicitly or implicitly coordinate them. I now elaborate the key relationships between these components.

Counting processes and formulas/expressions. The relationship between counting processes and expressions/formulas is shaded below in Figure 2. Note that as the arrow is bidirectional, I discuss both directions of this relationship.

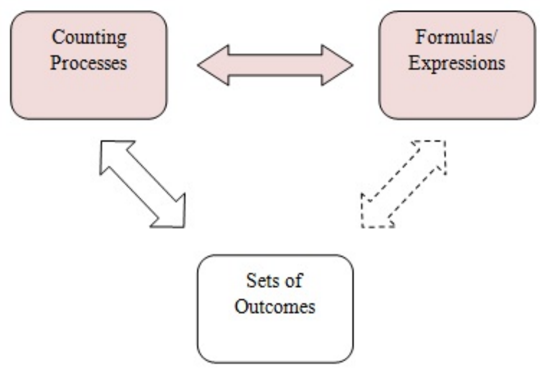


Figure 2: The Relationship between Counting Processes and Formulas/Expressions

I claim that, in the context of a counting problem, a given mathematical expression can often naturally be associated with a counting process. I want to be clear that in the discussion below, I am interested in students' constructions of the relationship between counting processes and formulas/expressions, and not the objective reality of this relationship, if there is one. The discussion is meant to elaborate this relationship that may arise for students in solving combinatorial enumeration problems, not to claim that there is a particular process that necessarily and universally underlies a given formula or expression.

Formulas/expressions → counting processes. A given formula/expression may elicit a counting process. The expression $\binom{5}{2}\cdot\binom{5}{3}$ is an example that highlights this direction of the relationship. This product of binomial coefficients can vary in what it represents. From one perspective, it simply represents a numerical value – we could calculate the product to arrive at 100. However, in the context of counting, this same product tends to signify a particular process. Specifically, it is an instance of the multiplication principle in which a typical element that is being counted is constructed in two stages. In the first stage, two objects are chosen from five distinct objects, and in the second three objects are chosen from five distinct objects; the multiplication indicates that the two stages are performed independently. I can further specify a context, such as a problem that states, "In how many ways can you form a committee of 2 men and 3 women, chosen from 5 men and 5 women?" In the context of such a problem, the expression can represent an even more specific process – choosing two of the five men, and then choosing three of the five women. Regardless of the context, however, counters can attribute combinatorial meaning to a mathematical expression in the form of a counting process.

Counting processes → formulas/expressions. In the opposite direction, we could conceptualize a counting process that generates an appropriate formula. If we wanted to count the number of ways of arranging 5 objects from a set of 10 distinct objects, there is a counting process that would allow us to do that, and this counting process could be expressed through an expression. We could consider the number of options for which object could go in the first position (10), then consider the number of options for the second position (9), etc., and using the multiplication principle we could arrive at an answer of $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$. There are thus formulas and mathematical expressions that can be generated by a particular counting process with which we might engage. In fact, this particular act of producing a formula from a counting process is often the end goal of solving a counting problem. In counting problems, often an expression or a formula (and not a numerical value) is the more desirable and meaningful solution to a problem; this can be particularly true of solutions with very large numerical answers.

Further comments on the relationship between counting processes and formulas/expressions. There may be more than one counting process associated with a single formula or expression, and there may be more than one formula associated with a given counting process. As an example of the former, we consider the expression $\binom{10}{5}$. This is a numerical expression with a numerical value; it is equivalent to $\frac{10!}{5!5!}$, or 252. If we consider the question "How many ways are there to choose a committee of 5 people from a faculty of size 10?," the answer is $\binom{10}{5} = 252$, but there are two different counting processes that could get us there, each represented by the same expression. We could first have arrived at the answer by choosing 5 of 10 people to be in the committee, yielding $\binom{10}{5}$. We also could have arrived at the answer by choosing 5 people *not* to be on the committee, also done in $\binom{10}{5}$ ways. So, while the expressions are the same in form, the processes by which we arrived at the expressions (and the reasoning behind each expression) differ.

It also may be the case that there could be more than one expression associated with some counting process. For instance, two students may have learned different expressions for the process of choosing a set of k objects from n distinct objects. For one student, an expression associated with that process may be $\binom{n}{k}$, for another it may be $\frac{n!}{(n-k)!k!}$. These are externally distinct expressions that may be associated with the same counting process.

There also may be distinct processes that arrive at different expressions, which accomplish the same counting result. For example, if we wanted to arrange 5 objects in 10 slots, we could use the multiplication principle to place objects successively in positions, arriving at $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$. This represents one process-to-expression pair. However, instead of directly arranging 5 of 10 objects in slots, we could first choose 5 of the 10 objects that we will arrange, done in $\binom{10}{5}$ ways, and then arrange them in 5! ways. This yields an answer of $\binom{10}{5}\cdot 5!$, which represents another process-to-expression pair. The expressions $10\cdot 9\cdot 8\cdot 7\cdot 6$ and $\binom{10}{5}\cdot 5!$ are equivalent, but they differ in form, and each represents a unique counting processes. Ultimately, the same end result is achieved (the number of ways of arranging 5 of 10 objects is calculated), but two different processes led to two different expressions.

There are many possible ways in which counting processes and formulas/expressions may interact. It may be beneficial for counters to be able to move back and forth between counting processes and formulas/expressions, and to recognize not only that a counting process can yield an expression or a formula, but that a given formula can also represent some counting processes. Being able to make sense of both directions of this relationship is an important aspect of evaluating alternative expressions.

Sets of outcomes and formulas/expressions.

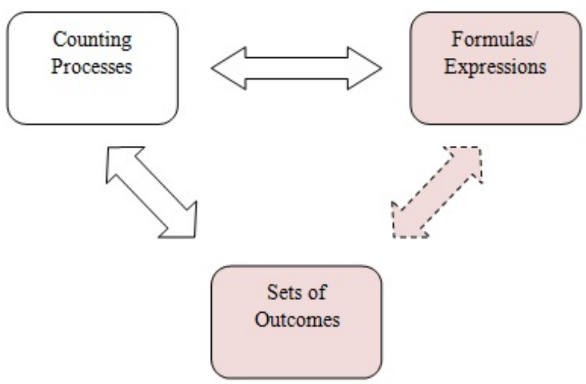


Figure 3: The Relationship between Sets of Outcomes and Formulas/Expressions

In the diagram of the model above (Figure 3), the arrow representing this relationship is dotted, because in the data this relationship was less clearly linked than the other two. I suspect perhaps for experienced counters there may be certain sets of outcomes that could be directly connected to particular formulas or expressions without having to consider a counting process. An example of this is an expression for a binomial coefficient, $\binom{n}{k}$. While there is an underlying counting process that it represents (choosing a subset of k objects from a set of n distinct objects), for some counters it may become an expression with encapsulated set-theoretic meaning. Specifically, it can be seen as the set of all possible k-element subsets whose elements come from some larger n-element set. I did not find evidence in the data that would help to flesh out the relationship, and as such it is a theoretical rather than empirical aspect of the model. I mention the relationship here primarily for the sake of completeness and to highlight it as a part of the model that could be examined more closely in subsequent research. I suspect that it may be the case that this particular relationship does not commonly arise directly, but rather that sets of outcomes and formulas/expressions tend to be connected *through* counting processes.

Counting processes and sets of outcomes.

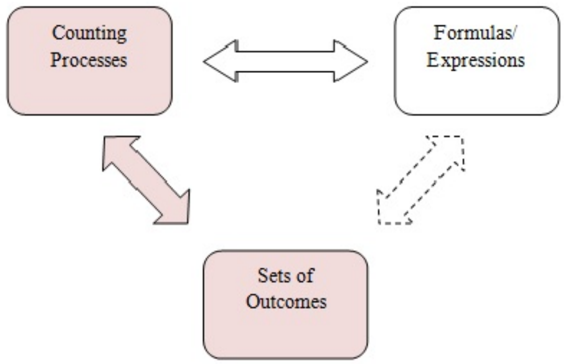


Figure 4: The Relationship between Counting Processes and Sets of Outcomes

As with the relationships between counting processes and formulas/expressions, the relationship shaded in Figure 4 below is bi-directional. Counting processes may generate some set of outcomes, and, conversely, a given set of outcomes may be enumerated (or its size may be determined) via some counting process. I elaborate the following example to describe the relationship between counting processes and sets of outcomes: "How many 3-letter 'words' are there using the letters A, B, and C (repetition allowed)?" The set of outcomes associated with this problem are the three letter words that satisfy the constraint, of which there are 27. There are multiple counting processes that could correctly answer the counting problem, and I discuss two such processes for this example. One possible counting process is first to apply the multiplication principle to consider the number of choices for the first letter, second letter, and third letter. The choices are independent, and, per the discussion of counting processes and formulas/expressions above, this process can be represented by the expression $3 \cdot 3 \cdot 3$, which gives an answer of 27. A second process breaks the problem into cases, organizing the words according to the number of distinct letters that appear in a particular outcome. The counting process involves enumerating each type of word and adding the cases. That is, we first consider the solution with only one type of letter (all A's, all B's, or all C's), then the solutions with exactly two types of letters (only A's and B's, only A's and C's, or only C's and B's), and finally the solutions with of all three letters. The three respective parts of the case breakdown have sizes 3, 18, and 6, respectively, which gives a total answer of $3+18+6=27$.

Counting processes → sets of outcomes. In this direction of the relationship, a counting process can be seen as generating some set of outcomes. Staying with the example of 3-letter words, the first process described above produces a particular listing of the set of outcomes. That is, by first considering that the first letter can be A, B, or C, and then noting that for each of those choices, the second letter can be A, B, or C, and so on, the set of outcomes can be generated. The tree diagram in Figure 5 makes the generation of outcomes more apparent; the structure of the diagram (the three points of branching) gets at the three-stage process of the multiplication principle, and the resulting list of the set of outcomes is to the right.

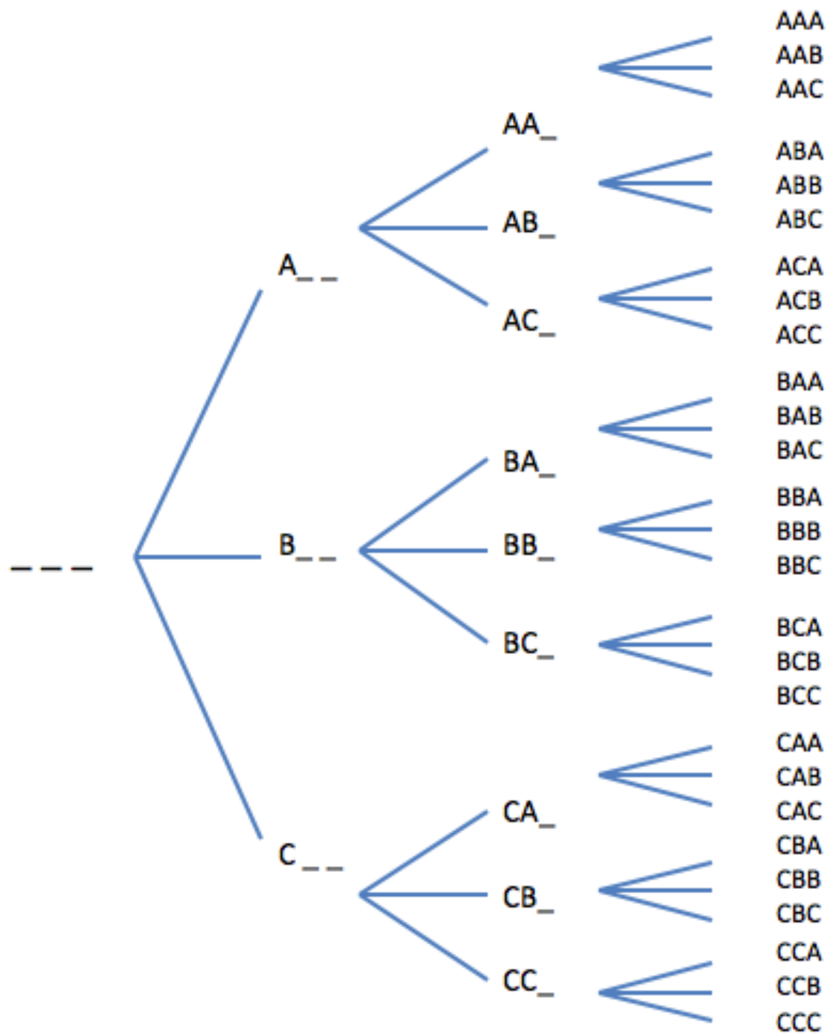


Figure 5: A Tree Diagram for 3-letter Words

In addition to generating a set of outcomes, a counting process can impose a structure onto a set of outcomes (and in fact different counting processes can result in different structures). In Figure 5, the counting process in the tree diagram actually organizes the set of outcomes into an alphabetical list, and, given the counting process of considering letter options for the respective positions, this makes sense. Alternatively, the second process of breaking the problem into cases and counting words based on the number of letters that appear can be seen as organizing the set of outcomes in a different way. In Figure 6 below, I have included on the left the alphabetical list that was generated by the multiplication principle process, and on the right organized list that was based on the number of repeated letters. The diagram below shows the two ways in which the different counting methods structured the set of solutions; there is a one-to-one correspondence between the set on the left and the set on the right. The set of outcomes is represented in both lists, and the cardinalities are the same, but the processes that yielded the set of outcomes differed. This example of two different counting processes illustrates the fact that a given counting process can impose a particular structure on the set of objects being counted.

AAA AAB AAC
ABA ABB ABC
ACA ACB ACC

BAA BAB BAC
BBA BBB BBC
BCA BCB BCC

CAA CAB CAC
CBA CBB CBC
CCA CCB CCC

AAA BBB CCC

AAB ABA BAA
BBA BAB ABB

AAC ACA CAA
CCA CAC ACC

BBC BCB CBB
BCC CBC BCC

ABC ACB
BAC BCA
CAB CBA

Figure 6: Two Ways to Structure the Set of Outcomes

Sets of outcomes → counting processes. The discussion above has focused on one direction of the relationship – how a student can generate (and organize) a set of outcomes from some counting process. I now discuss the other direction, in which a student can arrive at a counting process from a set of outcomes. For both of the processes in the 3-letter word example above, we could also think about starting with the set of outcomes, decide to organize that set in a particular way, and then come up with a formula to enumerate the set that is consistent with that specific organization of the set. For instance, we could have decided to start the problem by imagining (or actually) listing the outcomes alphabetically. This could have led to the consideration of choices for each letter, and ultimately to the process of implementing the multiplication principle. Or, we could have realized that the set of outcomes could be partitioned according to the number of distinct letters in each word, and we might have decided that we wanted to break up the outcomes accordingly. We could have then implemented a counting process that determined how many passwords were in each possibility then added to find the total. It is noteworthy that such consideration of the set of outcomes may be possible regardless of whether a student can conceptualize every element of the set[2].

Further comments on the relationship between counting processes and sets of outcomes. I propose that is often likely a back and forth association between counting processes and sets of outcomes. That is, a student could start a counting problem by making an initial attempt at a correct solution – he or she might choose a particular counting process that generates a set of outcomes. The student could then consider that set of outcomes and evaluate whether that set correctly answers the counting problem. In a case in which the generated set of outcomes does not align with the desired set of outcomes, the student might compare those sets and then return to the counting process to try to engineer a process that correctly enumerates the desirable set of outcomes. Such activity would involve movement back and forth between counting processes and sets of outcomes. In light of such activity, I contend that the link between counting processes and sets of outcomes can be (and should be) a very flexible, fluid relationship, in which students

[2] While the set of outcomes may exist as the set whose cardinality is the answer to a counting problem, a student may not necessarily want to (or be able to) consider the entire set of outcomes. Students might think about sets of outcomes conceptually (rather than concretely) or may theoretically organize sets of outcomes even when cardinalities are very large.

easily move from one component to another. Counting can be seen as an activity that relates counting processes to an underlying set of outcomes, and a set of outcomes can provide a way for students to ground their combinatorial activity, which can ultimately help to determine whether a counting process is correct. Part of what makes counting fascinating is that counting processes can seem logically sound, but they can yield incorrect answers. When this happens, it can be difficult to determine *why* a given counting process is incorrect. Students can gain much traction in determining whether a process is counting correctly by grounding their work in a set of outcomes.

Examples from the Data

In this section I use the model to talk through specific examples from the study from which the model emerged. These examples should serve to elaborate the components and relationships within the model in further detail, but even more they should demonstrate the model "in action." I have spent the preceding pages outlining details of the model, and now I illustrate how the model was used to analyze data from a study in which post-secondary students solved counting problems. In so doing, I hope to demonstrate the utility of the model. For the sake of efficiency, below I state four counting problems to which I will refer. Due to space I do not give detailed descriptions of the solutions (which are not necessary for the discussion); see Lockwood (2011) for further exploration of the problems.

Passwords problem: A password consists of 8 upper-case letters (repetition allowed). How many such 8-letter passwords contain at least 3 E's?

Cards problem: 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?

Groups of Students problem: In how many ways can you split a class of 20 into 4 groups of 5?

Test Questions problem: Suppose an exam consists of 10 questions, and you must choose 5 questions to answer. In how many ways can you choose 5 questions to answer if you must answer at least 2 of the first 5 questions?

In this section, I demonstrate three contrasting cases, each of which highlights one particular aspect of counting that students thought about in their work on the tasks. I examine students' work with cases, students' determination of whether or not order mattered, and students' recognition (and correction) of an instance of overcounting. In each of these cases, I use the model to frame the students' thinking and activity, comparing instances in which students did not draw upon sets of outcomes (Figure 7a) with instances in which they did draw upon sets of outcomes (Figure 7b).

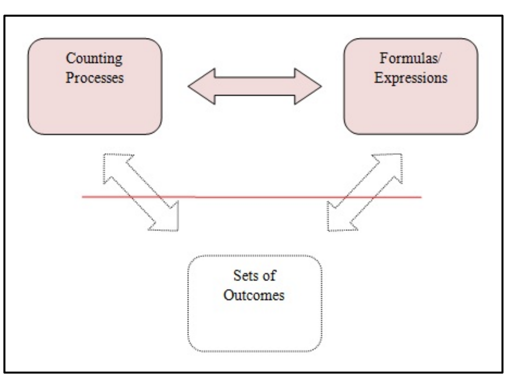


Figure 7a

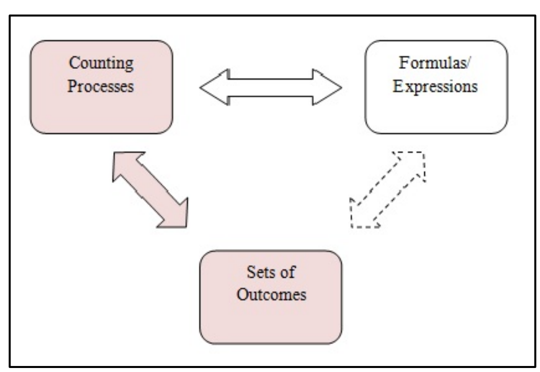


Figure 7b

Examples from the Data – Students Use Cases

In Kristin's work on the Passwords problem, she had used a case breakdown – to count the number of 8-letter passwords with *at least* 8 E's, she counted the number of passwords that contained exactly 3, 4, 5, 6, 7, or 8 E's. While she did not count each case correctly, she did sum her cases in her final answer. As we discussed her work, we had the following exchange.

E: Can I ask what made you think to do cases?

K: 'Cause it says at least 3, so I know I can have up to 8…

E: Can you say more about why you added?

K: I added them because … when it says "or" I always think of add. And for "and" I always think multiply…So cases I always know, add them.

Kristin's reasoning for why cases work is based on the keyword "or," which is a surface feature of the problem, and her consideration of cases seems to be made independently of considering the set of outcomes. I thus characterize her work on this problem by the relationships highlighted in the model in Figure 7a above.

In contrast to Kristin's work on the Password's problem, we consider Casey's work on the Cards problem. The correct answer to this problem is $9 \cdot 13 + 3 \cdot 12$, based on whether or not the first card (a face card) is also a heart. Casey got the problem correct, and when he explained his use of cases in the excerpt below, he seemed clearly to tie his work with cases with the set of outcomes.

E: How did you know then to go to the $9 \cdot 13 + 3 \cdot 12$?

C: I figured you could break it up into two different cases, because … you have a certain number of possibilities in this set of outcomes, and a certain number of possibilities in this set, and you just add them together and that would be the total number.

Unlike Kristin, Casey's thinking about cases is grounded in the set of outcomes; for him adding the cases makes sense not because of a key word in the problem, or because of a memorized trick, but because of an understanding that the set of outcomes can be partitioned into disjoint subsets. I would characterize Casey's thinking as being highlighted by the relationship in Figure 7b.

Examples from the Data – Students Consider Whether or Not Order Matters

As another set of contrasting cases, we look at students' determination of whether or not order matters in a problem, which is a common constraint in counting problems. Here I use the model to frame two different approaches to order. While working on the Passwords problem, Kristin had decided that she would use combinations for one aspect of the problem. She noted that she used combinations because she didn't "want order to matter."

K: I'm doing the combination ones because I'm pretty sure order doesn't matter with combination…I don't want order to matter.

E: Okay, and how come?

K: I'm not sure about that one (laughs). I just kind of go off my gut for it, on the ones that don't specifically say order matters or it doesn't matter.

When having to decide whether order matters, Kristin went "off her gut." And, even more, unlike the student in the next example, her notion of order mattering is not grounded in the set of outcomes. This is indicative of a fairly common phenomenon that can arise for students when considering order – if it is not clearly stated, students find it difficult to make a decision about order in a counting problem. Her thinking here could be highlighted in the model by Figure 7a.

In contrast to this, we highlight another students' work that does reflect a consideration of the set of outcomes (and would be highlighted by the relationship in Figure 7b). In Zach's work on the Groups of Students problem, he relates the decision about whether order matters to particular outcomes, specifying two particular outcomes in order to explain his reasoning about order.

Z: The order of the groups doesn't matter.

E: Okay, and what do you mean by that?

Z: A Group 1 with ACDEG, this is not distinct from GCDEA, where I swap place of any two students.

Zach recognized that he did not want order to matter in that part of the problem because two different outcomes, ACDEG and GCDEA, were not actually distinct from each other. This reasoning allowed Zach to correctly determine that order did not matter, and it also provided him with means by which to explain his thinking. This type of reasoning about order (through appealing to outcomes) has the potential to help students make meaningful decisions about order, rather than simply guessing based on their intuition.

Examples from the Data – Students Consider an Overcount

For this discussion, I offer contrasting cases in the same student's work. I use the model to highlight two aspects of the same student's work on the Test Questions problem; this emphasizes the way in which a student thought about the issue of overcounting. Marcus had initially arrived at an answer that is incorrect and that reflects a common error, $\binom{5}{2}\binom{8}{3}$. He explains his work in the excerpt below.

M: So, there's 5 choose 2 combinations of questions I can answer out of the first 5…So I have a remaining 8 to go, I have to choose 3 of them, three of those 8 to finish the 5 questions. It's pretty convincing.

We see in the excerpt above that Marcus argued through the counting process and arrived at the expression. In this initial work, though, there is no consideration of the set of outcomes, and I would thus characterize this part of his thinking as being represented by the diagram in Figure 7a. We see from his language above that Marcus seems to be convinced by the counting process that he utilized.

Some time later in the interview, as per the protocol, we revisited the problem. I presented Marcus with the correct solution of $\binom{5}{2}\binom{5}{3}+\binom{5}{3}\binom{5}{2}+\binom{5}{4}\binom{5}{1}+\binom{5}{5}\binom{5}{0}$, and he made sense of it, recognizing the case breakdown and noting that each case counted the number of ways two exactly two, three, four, or five of the first five questions, respectively. He thus found himself in the situation of comparing two expressions – his original solution, and the new expression that I had presented. Since it is not immediately clear whether these expressions are equivalent numerically, he used a calculator and realized there was a discrepancy between the two values – his incorrect answer is 560, and the correct answer is 226 – a difference of 334. Faced with this difference, Marcus realized that he might have overcounted.

M: So, just from what I have here, either they missed a bunch here, or I overcounted a bunch there. And I'm thinking I overcounted a bunch.

Marcus took time to think about how he might make sense of and explain the discrepancy. In Figure 8 below, we see evidence that he utilized of the set of outcomes to determine why the overcount had occurred. Specifically, we see that he appealed to a particular outcome (namely the set of questions {1, 2, 3, 8, 10}) that had been overcounted. To explain the overcount, Marcus drew the diagram in Figure 8 below. The rows of dashes in the diagram can be thought of as ten questions, and in the first row of the diagram he put x's in spots 1, 2, 3, 8, and 10, representing choosing those questions to solve. He noted that in the $\binom{5}{2}$ step, he could have chosen questions 1 and 2 (which he circled), and then in the $\binom{8}{3}$ step, chose 3, 8, and 10. Marcus drew the first row below and said the following:

> ***S***: So, I could have picked this one and this one [puts x's in and circled spots 1 and 3 in the second row]. Let's say, these guys. So now let's say, of the remaining 8 I pick 3 again, well that could be this one, this one, and this one [puts x's in 2, 8, and 10 in the second row]. So that's why I re-counted.

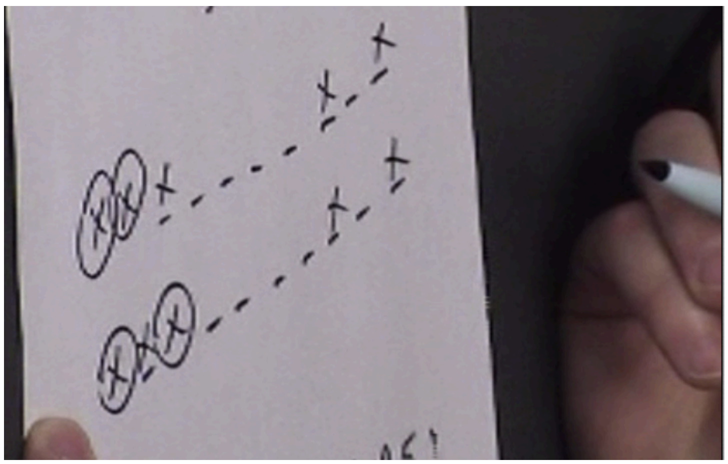


Figure 8: Marcus Uses a Particular Outcome to Explain an Overcount

So, Marcus pointed out that in the $\binom{5}{2}$ step, he could have chosen questions 1 and 3, and then in the $\binom{8}{3}$ step he could have chose 2, 8, and 10. Thus, the set {1, 2, 3, 8, 10} was counted more than once, and this explained to him why his original answer was too big.

What we see from Marcus, then, is that he was able to utilize the set of outcomes (and specifically a particular outcome) in order to address an issue of overcounting. This could be represented by the relationship in Figure 7b, in which he was able to relate his counting process with the set of outcomes, accounting for how the process was generating outcomes (and, in this case, generating some of them too many times).

The examples presented in this section are designed to show how the model can be used to frame different ways in which students might think about and approach various aspects of counting problems. These examples also highlight that there is potential benefit in focusing on sets of outcomes, and that it could be beneficial for students to consider the relationship between counting processes and sets of outcomes. Although the examples above emphasize the value of sets of outcomes, it is possible for the relationship between counting processes and formulas/expressions to be useful and correct; the utilization of formulas allows for much power and efficiency as we work. What I emphasize in these examples, though, and what I think the model can help me articulate, is that there is great potential in focusing on sets of outcomes. Indeed, there may be instances in which sets of outcomes are an indispensable part of understanding what is going on in some counting situations. As such, I suggest that attending to sets of outcomes should be seen as an intrinsic component of the activity of counting.

Conclusion

As discussed above, domain-specific models of student thinking about counting problems do not currently exist. The model presented in this paper offers a first attempt at addressing which kinds of concepts might be underlying students' combinatorial thinking, and in doing so it addresses a gap in the mathematics education research in the area of combinatorics. In addition to the overall potential of the model, I suggest that the model is innovative in emphasizing the significance and role of sets of outcomes. Data from the study discussed in this paper (described in detail in Lockwood, 2011) suggest that utilizing sets of outcomes can be particularly fruitful, and thus the model's emphasis on this aspect of counting is something that could be used effectively by mathematics education researchers.

Researchers could draw upon the components and the relationships in the model as a lens through which to describe and analyze students' counting activity. By highlighting relevant phenomena related to students' combinatorial thinking (and by facilitating the common articulation of such phenomena), the model may assist researchers in developing their understanding of students' conceptualizations of combinatorial ideas. Additionally, by getting a better sense of what aspects of counting students think about, understand, and struggle with, researchers may be more poised to conduct experiments to facilitate the improvement of teaching and learning related to combinatorics. While my study examined undergraduate students, I suspect that the components of the model may extend to K-12 student populations as well, and the model could serve as a tool for researchers at any level of investigation related to combinatorics education. There are also some potential implications for teachers and researchers implementing instructional design. In particular, I suggest that perhaps the components and the relationships can be something of which teachers are aware, perhaps considering ways to foster the relationship between counting processes and sets of outcomes in the classroom. This could come about by encouraging students to explicitly consider how a given counting process organizes the set of outcomes. Practical ideas could include encouraging students to engage in systematic listing (perhaps by utilizing tree diagrams or computer programs) and to have students clearly express what they are trying to count.

In sum, the model elaborated in this paper is meant to put forth an initial attempt at characterizing students' combinatorial thinking, providing ideas and common language that researchers can utilize in evaluating their own students' thinking and activity. While the model can certainly be further developed and investigated, by presenting the model I hope to offer the mathematics education community a starting point for the deeper investigation of students' combinatorial thinking.

References

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.

Hadar, N., & Hadass, R. (1981). The road to solve combinatorial problems is strewn with pitfalls. *Educational Studies in Mathematics, 12*, 435-443.

Kenney, M. J., & Hirsch, C. R. (Eds.). (1991). *Discrete mathematics across the curriculum, K-12: 1991 Yearbook.* Reston, VA: National Council of Teachers of Mathematics.

Lesh, R. & Doerr, H. M. (2000). Symbolizing, communicating, and mathematizing: Key components of models and modeling. In P. Cobb, E. Yackel, & K. McClain (Eds.). *Perpsectives on discourse, tools, and instructional design* (pp. 361-384). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lockwood, E. (2011). *Student approaches to combinatorial enumeration: The role of set-oriented thinking.* Unpublished doctoral dissertation, Portland State University – Oregon.

Maher, C. A. & Speiser, R. (1997). How far can you go with block towers? *Journal of Mathematical Behavior, 16*(2), 125-132).

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.

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

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: W. W. Norton & Company, Inc.

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* (pp. 267-307). Hillsdale, NJ: Erlbaum.

Strauss, A. & Corbin, J. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.

Thompson, P. W. (2008). Conceptual analysis of mathematical ideas: Some spadework at the foundations of mathematics education. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sepulveda (Eds.), *Plenary paper presented at the annual meeting of the international group of the psychology of mathematics education*, (Vol 1, pp. 45-64). Morelia, Mexico: PME.

Tucker, A. (2002). *Applied combinatorics* (4th ed.). New York: John Wiley & Sons.

Von Glasersveld, E. (1995). *Radical constructivism: A way of knowing and learning.* London: Falmer Press.

THE UNIT CIRCLE AND UNIT CONVERSIONS

Kevin C. Moore
University of Georgia
kvcmoore@uga.edu

Kevin LaForest
University of Georgia
laforesk@uga.edu

Hee Jung Kim
University of Georgia
hjk@uga.edu

While the unit circle is a central concept of trigonometry, students' and teachers' understandings of trigonometric functions typically lack connections to the unit circle. In the present work, we discuss a teaching experiment involving two pre-service secondary teachers that sought to characterize and produce shifts in their unit circle notions. Initially, both students experienced difficulty when given a circle that did not have a stated radius of one. The students relied on memorized procedures, including "unit-cancellation," to relate the unit circle to given circles. In an attempt to foster more robust connections between novel circle contexts and the unit circle, we implemented tasks designed to foster thinking about a circle's radius as a unit of measure. We report on the students' progress during these tasks.

Key words: Unit Circle, Trigonometry, Pre-service Secondary Teachers, Teaching Experiment, Quantitative Reasoning

Pre-service and in-service teachers often hold limited and fragmented understandings of central trigonometry concepts (Akkoc, 2008; Fi, 2006; Thompson, Carlson, & Silverman, 2007; Topçu, Kertil, Akkoç, Yilmaz, & Önder, 2006). Given that teachers have shallow understandings of trigonometry, it should come as no surprise that students construct disconnected understandings of trigonometric functions and topics foundational to trigonometry (Brown, 2006; Weber, 2005). In the hopes of improving trigonometry instruction, recent efforts (Moore, 2010, submitted; Thompson et al., 2007; Weber, 2005) have identified reasoning abilities and understandings that support robust notions (e.g., understandings containing flexible connections between trigonometry contexts) of the sine and cosine functions. Collectively, these studies highlight that attention must be given to students' and teachers' understandings of the things – angles, angle measures, the unit circle, right triangles, etc. – that trigonometry is about when attempting to improve their trigonometric understandings.

The present study explores two pre-service secondary teachers' (referred to as students from this point forward) thinking during an instructional sequence on trigonometric functions. We extend previous work in this area (Moore, 2010, submitted) by describing the students' evolving notions of the unit circle. Our findings illustrate relationships between students' measurement notions and their understanding of the unit circle. For instance, the students initially described the unit circle as a circle with "a radius of one," where "one" did not represent a measure in radii (e.g., one radius length). The lack of connection between the unit circle and measuring in radii inhibited their ability to use the unit circle in novel contexts and provide quantitative meanings to their calculations. We also discuss the deep-rooted nature of the ways of thinking the students held upon entering the study, including a reliance on dimensional analysis, which provides broader implications for mathematics education and the preparation of K-12 teachers.

Background

Research (Akkoc, 2008; Fi, 2006; Thompson et al., 2007; Topçu et al., 2006) on pre-service and in-service teachers' trigonometric knowledge suggests that teachers are lacking the content knowledge necessary to support their students' learning of trigonometry. Teachers are often tied to discussing trigonometric functions in a right triangle context while making only superficial connections to circle contexts (Akkoc, 2008; Thompson et al., 2007; Topçu et al., 2006). Complicating the issue, several studies maintain that teachers hold disconnected and shallow understandings of angle measure that influence their ability to create connections between trigonometry contexts. For instance, when synthesizing the findings of a sequence of related studies (Akkoc, 2008; Topçu et al., 2006), the authors suggested that teachers' reliance on degree angle measure likely restricts their understandings of trigonometric functions to triangle contexts. Compatible with teachers who entrench trigonometric functions in right triangles, students' (some of who will become teachers) notions of trigonometric functions are restricted to right triangle contexts (Weber, 2005) and they encounter difficulty reasoning about trigonometric functions in a circle context (Brown, 2005).

While teachers and students frequently lack robust understandings of trigonometry, recent studies (Moore, 2010, submitted; Weber, 2005) have made progress in identifying meanings and reasoning abilities that support connected understandings of trigonometric functions. When comparing the progress of students enrolled in a traditional trigonometry course to that of students enrolled in an experimental trigonometry course informed by Gray and Tall's (1994) theoretical notion of precept, Weber (2005) noted a difference in the two groups' abilities to leverage the geometric objects of trigonometry (e.g., the unit circle). Students in the traditional group were unable to productively use these objects on their own accord, while students in the experimental group used the unit circle to solve novel problems. In light of his findings, Weber argued future investigations should explore how to support students in conceptualizing the geometric objects of trigonometry in ways that posit students to use these objects as tools of reasoning in novel situations.

Following Weber's suggestion, as well as several calls (Bressoud, 2010; Thompson, 2008) for revising trigonometry (and angle measure) instruction, Moore investigated precalculus students' angle measure conceptions (submitted) and the role of angle measure in students' construction of the sine function (2010). These studies illustrated that arc length images of angle measure, in combination with reasoning about the radius as a unit of measure, can create foundations for coherence between the trigonometry contexts. Additionally, the students' thinking highlighted the important role of quantitative reasoning in constructing connected understandings of trigonometric functions and angle measure.

Quantitative Reasoning and Measurement

Quantitative reasoning (Smith III & Thompson, 2008; Thompson, 1989) provides a model of how reasoning about quantities (e.g., measurable attributes of objects) and relationships between these quantities (e.g., a multiplicative comparison between two quantities) can form a foundation for student learning and the emergence of meaningful mathematical formalisms (e.g., graphs and formulas). A central premise of quantitative reasoning is that quantities exist in the mind and are thus unique to the individual. Research has illustrated that students' conceptions of quantities and quantitative relationships should not be taken as a given or considered trivial (Moore & Carlson, 2012; Smith III & Thompson, 2008; Thompson, 2011).

Measurement is a critical aspect of quantitative reasoning, with Thompson (2011) attributing Steffe (1991a, 1991b) with a foundational analysis of the mental operations that generate quantity. In Thompson's (2011) telling of quantitative reasoning, he suggested that an oft overlooked and nontrivial aspect of measurement is reasoning about magnitudes. To explain magnitude, Thompson alluded to Wildi's (1991) description of magnitudes, which is based on the notion that the magnitude of a quantity is not dependent on the unit used to measure the quantity. For instance, one's height at any given moment in time is the same magnitude regardless of the unit used to measure the height. Given that the measure of a quantity is *a* units, the magnitude of a quantity is *a* times as large as the magnitude of the unit used to make the measure.

Thompson (2011) argued that magnitude reasoning is critical for a deep understanding of quantity and he cautions that it is necessary for schooling to attend to this reasoning on a repeated basis. Consequently, Thompson suggested a unit conversion approach that centers on reasoning about the relationship between the measure of a quantity and the magnitude of the unit used to make the measure for various unit magnitudes as one setting for repeated experiences in magnitude reasoning. To borrow an example from Thompson, "if the measure of a quantity is M_u in units of u, then its measure is $12M_u$ in units of magnitude $(1/12)\|u\|$ and its measure is $(1/12)M_u$ in units of magnitude of $12\|u\|$" (2011, p. 21).

This *magnitude-measure* approach to unit conversion, in combination with changing the unit magnitude to the radius, provides a way to think about every circle as the unit circle. If the radius of a circle is 4.2 feet, one can reason that the radius has a magnitude that is 4.2 times as large as the magnitude of a foot. If the radius is now thought of as the unit magnitude, it follows that measures in radii will be 1/4.2 times as large as corresponding measures in feet. Hence, to convert a measure in feet to a measure in radii, one divides by 4.2 (or multiplies by 1/4.2). This line of reasoning provides a natural way to (a) address the output of trigonometric functions as ratios, (b) give meaning to said ratios (e.g., measuring in radii), and (c) provide a conceptualization of the unit circle that can be applied to a circle whose radius length is given in any unit other than radii (Figure 1).

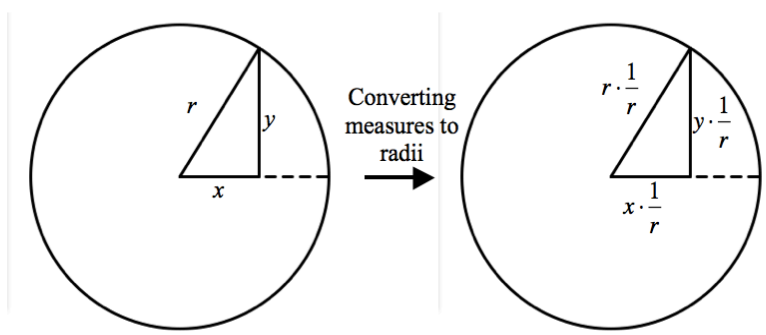


Figure 1 – Unit Circle, Ratios, and Units of Measure

In comparison to a magnitude-measure approach to unit conversion, a more common approach found in mathematics, engineering, and physics courses is that of dimensional analysis. Students sometimes refer to this approach as *unit-cancellation*. Dimensional analysis typically consists of starting with a measure (4.5 feet), identifying two equivalent measures with one measure in the given unit and the other measure in the desired unit, and then cancelling units to decide what ratio to multiply by the given measure. For instance, converting a measure of 4.5 feet to a number of centimeters would follow:

- I have a measure in feet and wish to find the equivalent measure in centimeters.
- There are 30.48 centimeters in 1 foot, or there are 0.0328 feet in 1 centimeter.

- The desired measure is $4.5\ \cancel{\text{feet}} \cdot \dfrac{30.48 \text{ centimeters}}{1\ \cancel{\text{foot}}} = 137.16 \text{ centimeters}$, or

  $4.5\ \cancel{\text{feet}} \cdot \dfrac{1 \text{ centimeter}}{0.0328\ \cancel{\text{foot}}} = 137.16 \text{ centimeters}$.

The calculations performed in dimensional analysis can appear identical to those resulting from reasoning about magnitudes (e.g., the magnitude of a centimeter is 1/30.48 times as large as the magnitude of a foot, and thus a quantity's measure in centimeters is 30.48 times as large as the quantity's measure in feet). But, dimensional analysis circumvents reasoning about magnitudes and avoids providing a quantitative reason for using the operations of multiplication or division. Dimensional analysis leaves the basis for the conversion calculations implicit, and instead treats units as if they are things that can be discarded through procedural rules.

On the basis of the argument made by Thompson (2011), and the complexity involved in constructing measurement schemes (Steffe, 1991a, 1991b), one shouldn't make the assumption that students understand the underlying mathematics structure when executing unit conversions, particularly when using dimensional analysis. In fact, there is evidence (Reed, 2006) that dimensional analysis can mask important mathematical ideas and lead to decreases in student performance. Reed (2006) predicted that students would realize improved performance at constructing equations for word problems upon completion of instruction on dimensional analysis. Instead, he found that student performance decreased and he partially attributed this decrease to students' attempts at memorizing a rote procedure. Specifically, two students claimed that dimensional analysis helped them to memorize equations and write formulas without understanding. The students' claims, combined with their actions, suggest that a focus on dimensional analysis did not support the development of understandings that help them solve word problems.

Methodology

Stemming from radical constructivist theories of knowing and learning (Glasersfeld, 1995), we consider each individual's knowledge fundamentally unknowable to any other individual. Reflecting this stance, we sought to build and test *models of the students' thinking* in an attempt to obtain viable models of the students' mathematics that explained their observable behaviors. We used qualitative methods to gain insights into their thinking when attempting to solve the proposed tasks. Specifically, we used a teaching experiment methodology (Steffe & Thompson, 2000) to pursue the research questions of:

- What are students' notions of the unit circle and how do these notions impact their use of trigonometric functions?
- What are the critical ways of reasoning involved in conceptualizing the unit circle in ways that support its use in novel situations?

Subjects and Setting

The study's participants (Bob and Mindy) were two undergraduates enrolled in a pre-service secondary mathematics education program at a large state university in the southeast United States. Both Bob and Mindy were second year students at the time of the study. We chose the students on a voluntary basis while they were enrolled in their first course in the education program. Bob and Mindy were the only students (out of 10) to volunteer for the study. Previous to the study, the content course (in which the first author was the instructor) covered angle measure and an introduction to the sine and cosine functions using tasks that were in line with those outlined in previous studies (Moore, 2010, submitted).

Data Collection and Analysis Methods

We used a teaching experiment methodology (Steffe & Thompson, 2000) to develop and test models of the students' thinking. Each student participated individually in five 60- to 90-minute teaching sessions that took place within a span of eighteen days. The predominant focus of the present study is on the first two sessions with each student (later sessions focused on the students' notions of periodicity). During their participation in the study, the two students did not attend the regular class sessions of the content course. The lead author acted as the instructor for each teaching session, with the second and third authors acting as observers. All three authors met between the teaching sessions to discuss their observations.

We used an open and axial coding approach (Strauss & Corbin, 1998) in combination with a conceptual analysis (Thompson, 2000, 2008) to analyze the data. More pointedly, we first sought to build viable models of each student's thinking over the course of the study. We then compared and contrasted each student's thinking in order to determine how his or her thinking evolved over the course of the study. For instance, we compared and contrasted Bob's notions of the unit circle over the course of the study in an attempt to document how his understanding of the unit circle evolved and what operations led to these evolutions. After conducting the same analysis of Mindy's notions of the unit circle, we juxtaposed the two students' progress in order to gain deeper insights into their thinking.

Results

In this section, we discuss the first session with each student to highlight his or her notions of the unit circle. Against this backdrop, we describe the design of the subsequent teaching sessions, the students' activity during these teaching sessions, and shifts in the students' notions of the unit circle. We draw explicit attention to ideas of measurement, including measure conversions, and how these ideas might have influenced the students' unit circle notions.

Session One

During the pre-interviews, both students frequently drew "the unit circle" when solving various tasks. As an example, when solving the Arc Length Problem (Figure 2), Bob drew a "unit circle" that was distinct from the given circles to justify his answer. His solution consisted of first determining that the given angle had a measure of 0.611 radians. He then multiplied each given radius length by 0.611 to (correctly) determine specified values for the arc lengths, claiming, "0.611 would be the number of radius lengths within, on the unit circle." When asked to explain further, Bob drew a separate circle that he called "the unit circle" and marked the radius as "1." At this time, Bob attempted to draw an angle with a measure of 0.611 radians on his new circle, but he had difficulty doing so despite drawing the unit circle directly beside the given circle and angle. He eventually marked an arc that he claimed was less than $\pi/4$ radians.

Given that the following angle measurement θ is 35 degrees, determine the length of each arc cut off by the angle. Consider the circles to have radius lengths of 2 inches, 2.4 inches, and 2.9 inches (*figure not to scale*).

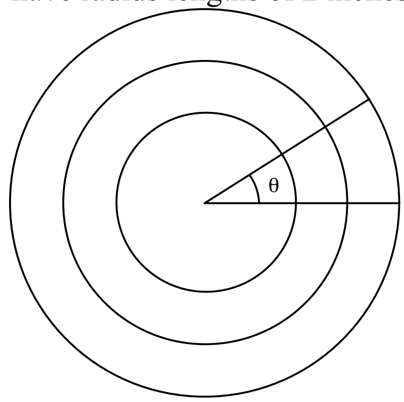


Figure 2 – Arc Length Problem

An important point to note about Bob's actions is that although he determined the correct arc lengths and executed a calculation that might imply reasoning about 0.611 as the number of radius lengths along each subtended arc length on the given circles, he only referred to 0.611 as the number of radius lengths along "the unit circle." We took Bob's actions of drawing a distinct unit circle and subsequent difficulty in identifying the angle that had a measure of 0.611 radians as suggestive of a disconnect between "the unit circle," the given circles, and the measure of 0.611 radians. More specifically, Bob seemed to understand that 0.611 radians conveys that 0.611 radius lengths lie along "the unit circle," but even after extended questioning it was not clear that Bob understood that each determined arc length is 0.611 radius lengths when measured with the corresponding circle's radius.

To test our assumption that Bob did not connect radian or radii measures to the given circles' radii and that he relied instead on a distinct "unit circle" to reason about these measures, we asked him to determine an angle measure when given a subtended arc length (1.2 inches) and the radius (3.1 inches) of that circle (Figure 3).[1] Bob began the problem by drawing a second circle to "convert to radians" (Excerpt 1).

Excerpt 1

Bob:	Problems like this, I think I usually convert to radians first. I think radians are easier to use than degrees. So I would divide this (*the radius*) by three point one to get one. So like if this were here (*drawing a new circle*), that's our angle again, copied, dividing by one gives that this is one radian (*labeling the radius as* 1). And I'll divide this (*the arc length*) by three point one also (*using calculator*).
Int.:	Ok, so why would you divide that by three point one also?
Bob:	So one point two divided by three point one is point three eight seven. And then I could say, well you wanna know theta (*labeling theta in the circle*), right. So I know there is a relationship between the arc length, s is equal to r times theta (*writing corresponding formula*). So our arc length here is point three eight seven, is equal to, our r is one, is equal to theta, and then you just multiply that by three point one to get the arc length in inches. Well this is in radians (*referring to point*

[1] We provided Bob the original diagram with the given measures.

three eight seven that came from the formula), and then if you multiply that by three point one (*using calculator*), you get, well obviously, ya, theta is one point two also, in inches, or, ya. Or I guess it would be degrees.

Consistent with Bob's actions on the previous problem, he first sought to convert measures to radians and drew a circle (which he later referred to as "the unit circle") distinct from the given circle to discuss the converted measures. After dividing each measure by the given radius, Bob used the formula $s = r\theta$ with r=1 and s=0.387 to determine θ in radians. This suggests that Bob did not interpret the quotient 1.2/3.1 or the number 0.387 as the measure of an arc in radii or radians until he used the formula to determine θ, adding further evidence to our claim that there was a disconnect between the given circles, "the unit circle," and measures in radians or radii. The division of the given measures by the radius length did not represent using the given radius length as a unit of measure for Bob, nor did the calculation yield an angle measure in radians.

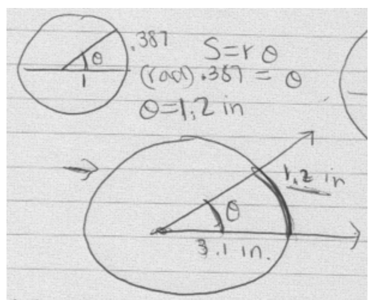


Figure 3 – Bob and the Unit Circle

We note that Bob ended the discussion in Excerpt 1 by claiming that θ is also 1.2 (after multiplying 0.387 times the given radius). He was unsure of the unit for this value, claiming that 1.2 represents either inches or degrees. Bob eventually concluded that 1.2 represents a number of inches due to the given 1.2 inches, but he was unable to give a justification for the calculation used to determine this value ($0.387 \cdot 3.1$). Bob had the same difficulty on the Arc Length Problem when he multiplied the determined radian measure by the given radius lengths, which is consistent with our hypothesis that radian measures did not imply a multiplicative relationship between the arc lengths on the *given* circles and the corresponding radius lengths.

Like Bob, Mindy only discussed the unit circle as a circle distinct from given circles. For instance, Mindy gave the following statement after suggesting that she needed to use the unit circle to give a meaning for radian angle measures (Excerpt 2).

Excerpt 2

Mindy: The unit circle is just, unit circle (*writing 'unit circle'*) equals circle with radius of one length, unit. You know, you could be talking about inches, you could be talking about kilometers. I mean, really it's just the unit is one of them. So in a lot of cases when you're talking about the unit circle, you're not talking about a designated unit. You're just talking about a radius of one. A circle with a radius of one (*writing this phrase by the term 'unit circle'*).

It is difficult to tell from Mindy's initial statements (lines 1-2) if she connected the unit circle to using the radius as a unit of measure, as Mindy did suggest that the unit circle is a circle with "a radius of one length." But, she then described that the "one" represents a common unit of length measure (e.g., inches or kilometers) or a unit-less number. This statement implied that the "one" she associated with the unit circle did not represent the radius as forming the unit of measure.

To gain further insights into Mindy's thinking, we subsequently asked her to determine an angle measure when given an arc length (6.6 inches) subtended by the angle and the radius (2.4 inches) of that circle. She first drew "the unit circle" (Figure 4) and then determined the angle measure (Excerpt 3).

Excerpt 3

Mindy: Maybe I can simplify this by creating a unit circle and converting these measurements to what they would be. So this is going to be our original circle (*writing this phrase by the given circle*), and then this is going to be the unit circle (*writing this phrase and drawing a new circle*). So, here's your center. So we know that, just, by nature the unit circle is going to have a radius of one, and because we're already given the unit, we can go ahead and say one inch (*writing 1 inch below the radius*). So if, um, so we can say, using ratios again, we're trying to find what the equivalent length of the arc length would be.

In this case, Mindy defined "the unit circle" to have a "radius of one…one inch" because the given measures represent a number of inches. Mindy then created an equation by matching values (1 inch to x inches and 2.4 inches to 6.6 inches) to obtain 2.75 inches for "the arc length in the unit circle." After creating an equation, and while performing calculations to determine x, Mindy focused on the units associated with each number. For instance, she claimed that her ratio was correct because the units matched in the original ratio and the answer had the correct units (2.75 *inches*). She then switched the unit to radians because the problem asked for an angle measure.

Mindy drew a second circle when solving the problem, which was consistent with our conjecture that Mindy did not connect the unit circle to using the given circle's radius as a unit of measure. The measures she labeled on "the unit circle" were not measures of the quantities on the given circle (e.g., Mindy conceived of two arc lengths, one on the given circle and one on the unit circle). Also, her actions did not suggest that she conceived of that circle's radius (or the given circle's radius) as a unit of measure, instead associating the length with the unit inch. In line with this claim, Mindy used an equation between two ratios that stemmed from correspondence thinking and comparing units to determine the unit circle arc length, as opposed to dividing the given arc length by the radius.

At this point, Mindy's and Bob's thinking appeared to be compatible. However, immediately after Mindy obtained the answer to the problem, she noted, "We see that, what we knew all along, we're just dividing this arc length by the radius, which would be putting it into terms of radians, I mean radii." We did not expect this statement and asked her why she didn't use this method from the beginning. She responded, "Once again units, just understanding. But if you know that you're trying to get the radius to be one, and…find out what this is (*pointing to original arc length*) in terms of the radii, we would be dividing by what the given radius is." Mindy's response implies that she did consider determining the given arc length in radii, but she considered comparing units as a solution representative of a better understanding.

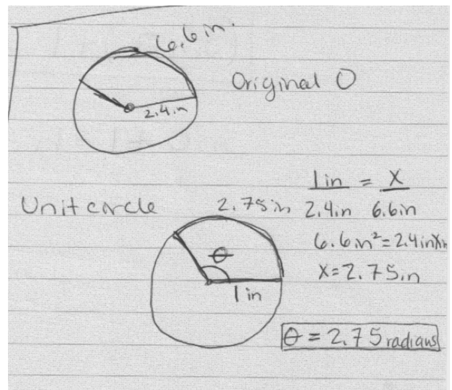


Figure 4 – Mindy and the Unit Circle

Collectively, Bob's and Mindy's actions during the pre-interview indicate that neither student conceptualized the unit circle as connected to using a given circle's radius as a unit of measure. We note that both students did divide given measures by the radius during the first session. For instance, Bob divided by the radius to relate the given radius to a "radius of one" on the unit circle. To him, dividing by radius measures did not yield measures in radii or radians, even when labeled on his "unit circle," and it was not until relating these numbers to the formula $s = r\theta$ that Bob determined a radian measure.

Mindy, on the other hand, did conceive of dividing measures by the radius as representing a number of radii. She reasoned that these values were equivalent to the corresponding radian measures on the unit circle. Like Bob, she characterized the unit circle as a circle with a "radius of one." However, her understanding that dividing by the radius represents measures in radii enabled her to more flexibly solve the given problems than Bob. Still, Mindy's unit circle was distinct from circles that did not have a radius with a given length of one, and she preferred giving measures on the unit circle a specified unit. She emphasized multiple times throughout the interview that she was most comfortable using dimensional analysis and she equated dimensional analysis to "understanding." We note that Mindy did not mention *radii* (or *radius lengths*) as a possible unit for dimensional analysis and treated the result of dividing two like unit values (e.g., inch measures for an arc length and radius length) as a unit-less value, despite describing this result as a number of radii. Her discomfort with this value was likely connected to her conviction that dimensional analysis not only verified the correctness of answers, but also implied understanding.

The students' ways of thinking about the unit circle enabled them, for the most part, to solve angle measure problems. But, their ways of thinking became problematic when attempting to use trigonometric functions in novel situations. When given problems that asked them to use trigonometric functions to relate angle measures and other quantities, they encountered difficulty moving between given circles and "the unit circle," with them often losing track of the meanings of obtained numbers and executed calculations. The students could identify that $\sin(0.5) \approx 0.48$ implies that the y-coordinate on the unit circle is 0.48 at an angle of measure 0.5 radians. Yet, to Bob and Mindy, the number 0.48 did not entail a unit of measure (including radians or radii). Thus, this number did not lend itself to the method of dimensional analysis, nor was it directly tied to using the radius of a given circle as a unit of measure.

Putting the Unit in the Unit Circle

In response to (a) the students' propensity to reason about the unit circle as distinct from given circles, (b) their focus on dimensional analysis when performing calculations, and (c) their absence of associating the unit circle to measuring in radii, we designed the second session activities to foster reasoning about the varying relationship between the measure of a quantity and the magnitude of the unit used to measure the quantity. One of our goals was to have the students come to view lengths as taking on multiple measures all at once. We conjectured that this reasoning would support their conceptualizing a circle's radius as simultaneously taking on measures in standard units (e.g., inches) and radii. We also speculated that such reasoning would support the students in coming to view the unit circle as tied to measuring in radii, with all circles having a *radius of one radius length* (versus a *radius of one*).[2]

As a first task, we gave the students[3] a picture of a stick and questions along the lines of:

1. What does it mean for the stick to have a length of 3.4 feet?
2. Given that there are 12 inches in one foot, how long is the stick when measured in inches? Given that there are 300 feet in a football field, how long is the stick when measured in football field lengths?
3. Given that a *fraggle* is a unit of measure that is 221 times as large as 2 feet, what is the length of the stick when measured in fraggles?
4. When answering the above questions, did the length of the stick change?

In our attempt to foster reasoning about how variations in the measuring unit influence the numerical result of measuring a quantity, we added the stipulation that the students were not to use formulas, dimensional analysis, or written expressions, unless absolutely necessary. These stipulations caused difficulties for both students. Mindy claimed, "Not getting to use unit-cancellation is confusing me, because that's how I'm doing them even when it's *not* (*using her fingers to make an air-quote gesture*) unit-cancellation."

As Bob and Mindy worked through the first few questions of the task, they used equivalent measures and imagined unit-cancellation to determine the conversions. For instance, to determine the length of the stick in inches, both students reasoned that 12 inches is equivalent to 1 foot, and subsequently multiplied the given measure by 12/1 (12 inches divided by 1 foot). To determine the measure of the stick in football fields, Bob and Mindy multiplied the given measure by 1/300 (1 football field divided by 300 feet), as 1 football field is equivalent to 300 feet. The students' solutions were not surprising in light of their focus during the first session.

It is important to note that we did not observe the students explicitly making comparisons between unit magnitudes when making conversions. The students did consider equivalent measures (e.g., 12 inches and 1 foot), but their actions did not imply that they compared the magnitudes of the units for these measures (e.g., the unit inch is 1/12 times as large as the unit foot). For instance, when prompted to explain their ratios (e.g., 1/300), they only described them as two "equal measures," as opposed to reasoning that the ratio conveys a multiplicative relationship between two unit magnitudes or two measures (Figure 5).

---

[2] We note that mathematicians and textbooks often define *radius* as a distance, thus the phrase *a radius of one radius length* might seem redundant. What we intend to point out is that the *one* in the phrase entails a unit – the radius. To students, *a radius of one* does not necessarily entail thinking about *one* as entailing the radius as a unit.

[3] Like the first session, we met with each student individually.

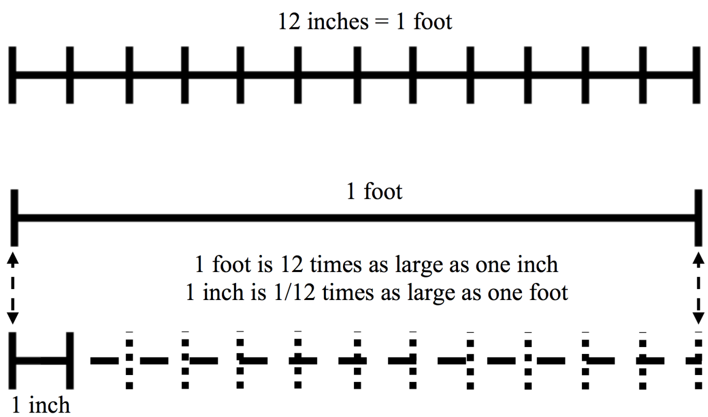


Figure 5 – Equivalent Measures vs. Comparing Magnitudes

Mindy's response to the third question listed above further confirmed a focus on equivalent measures without explicit attention to comparing unit magnitudes. She first decided to determine how many fraggles are in one foot (Excerpt 4).

Excerpt 4

Mindy:	I'm going to compare my fraggles to, I guess feet, so I want to make a conversion from feet to fraggles. So I need to know what the conversion is from one foot, not two feet. I'm just going to divide two twenty one by two. So now I know that there's gonna be, for every fraggle it's one-hundred and ten point five feet (*writing '1/110.5'*). Does that make sense? So, I'm just going to be dividing three point four by the number of feet in a fraggle. So now I know that there are point zero three zero seven seven fraggles here.
Int.:	So could you explain that to me a little bit, how you determined that?
Mindy:	Ya, I felt that the, I guess the relationship given, fraggles to feet isn't very useful because I want to know one foot. Because that's what I'm given, the stick in feet. So that's the first thing I did. I need to know the conversion from feet to fraggles, and it needs to be from one foot to a certain number of fraggles. So that's the first thing I did.
Int.:	So could you tell me how you determined that, the one foot to fraggles.
Mindy:	I just, I know that I want it to be one foot, so technically I divided both of these by two to get one foot. And then I just did as I did like before. Multiply out to get the stick in fraggles.

Consistent with her previous actions, Mindy first sought to determine the fraggles measure that was equivalent to one foot, and as a result assimilated the 221 in the problem statement as the number of fraggles equivalent to 2 feet. After determining the number of fraggles equal in length to one foot, she used these equivalent measures to then determine the length of the stick measured in fraggles. After directing her to reread the problem statement and describe the meaning of each value, she maintained that 221 fraggles were equivalent to 2 feet and that 110.5 fraggles are in one foot.

In response to Mindy's focus on equivalent measures without attending to comparisons between unit magnitudes, we asked her to describe key components of measurement. She responded, "We have to start off with an object of a particular length…this object is handy for comparing other objects to that one object…it's a comparison between two things." In response to Mindy raising the idea of a unit magnitude (e.g., "an object of a particular length"), we directed her back to the second question and asked her how the *units* compared. She claimed that an inch was 1/12 times as large as a foot. After she identified that this corresponds to the measure in inches being 12 times as large as the measure in feet, we asked her if the measure in football fields can be determined in a similar manner. Mindy stated that the unit is 300 times as large, and thus the measure should be "way small, a three hundredth."

With Mindy appearing to coordinate unit magnitudes and measures, we asked her to return to the fraggles question. She first restated that there were 110.5 fraggles in a foot. We then asked her to only discuss how the unit magnitudes compare, as opposed to determining equivalent measures, and she responded, "Oh noooo (*laughing*). It's the opposite of what I did. Right, ok." At this time Mindy realized that the magnitude of a fraggle was 221 times as large as 2 feet, or 442 times as large as one foot, leading to measures in fraggles being 1/442 times as large as measures in feet.

Over the course of the stick task, and a subsequent task that asked the students similar questions about attributes of a circle, we observed both students coordinating quantities' measures and the magnitude of the unit used to make these measures. At first, they had difficulty separating comparisons between measures and comparisons between unit magnitudes, often assuming the two were the same (e.g., if the unit magnitude is halved, the measure is halved). Both students claimed that reasoning about magnitude-measure was unnatural to them and they could not recall experiencing such reasoning previous to our instruction. When discussing unit magnitudes with Bob, he claimed, "I feel like the measurement itself changes in units, but it's the same measurement regardless, if that makes any sense…It's like when you measure things in radians and degrees, the angle is still the same." He gave this description when discussing that an arc length has two measures (e.g., radii and feet) that both represent the same magnitude.

Upon the completion of the two tasks, and at a point that we conjectured both students had used magnitude-measure reasoning to identify measures in radii on a circle (including that the radius has a measure of 1 radii), we presented the students with the "Which Circle?" problem (Figure 6).

Consider circles with a radius of 3 feet, 2.1 meters, 1 light-year, 1 football field, and 42 miles. Which, if any, of these circles is a unit circle?

Figure 6 – Which Circle?

Bob responded to "Which Circle?" by stating, "I guess, in retrospect, or in theory, they could all be unit circles just by dividing by their corresponding lengths." He then described dividing each given radius by that radius to obtain a radius of "one" and claimed, "You're scaling [the radius] down." This response suggested that Bob imagined the given circle being scaled down to another circle with a radius of "one." However, he then paused for several seconds and stated, "Well, not really scaling it down. I'm really just changing the unit again." By connecting the task to the previous problems, Bob concluded that each circle could be thought of as the unit circle if that circle's radius is made the unit of measure.

Mindy provided the following response to "Which Circle?" (Excerpt 5).

Excerpt 5

Mindy: The unit circle doesn't even require a specific unit other than the radius. I guess that's why it's called the unit circle is like the radius is always just one unit… If we made three feet our radius we would just think about the circle in terms of radii instead of feet then it would be a unit circle. Every circle has a radius, so if you just want to talk about the circle in terms of that unit the radius then every circle is a unit circle…as long as you are considering that the radius is one unit (*holding her hands apart to signify a length*). Like perhaps it's not one unit, I mean it's not one meter in length, but it's one radius in length.

In this example, Mindy described the unit circle as directly tied to using the radius as a unit of measure. Differing from her previous actions, Mindy also noted that every circle can be thought of as the unit circle, regardless of the given radius measure. Both students' explanations of the unit circle suggest an understanding that each circle's radius could simultaneously be measured in the given unit or in a unit equivalent to the radius length.

In light of the students' answers to "Which Circle?" we expected that the students' ways of thinking about magnitudes and measures might support them in using the unit circle in fundamentally different ways than the first teaching session. To test this conjecture, we asked both students to provide a solution to the following problem (Figure 7).

> An arctic village maintains a circular cross-country ski trail that has a radius of 2.5 kilometers. A skier started skiing from position (2.4136, 0.6513), measured in kilometers, and skied counter-clockwise for 13.09 kilometers where he paused for a brief rest. Determine the ordered pair (in both kilometers and radii) on the coordinate axes that identifies the location where the skier rested.

Figure 7 – Ski Problem

Bob responded to this problem in a way that implied he was thinking about the unit circle in ways that differed from his thinking during the first session. For instance, to begin the problem, he stated, "I guess what I'd want to do is scale [the unit] down again by a factor of… well, 1 radian is 2.5 kilometers. So by changing it we're going to multiply the values of the problem by a factor of 1 over 2.5." Bob followed this by drawing a new circle, an action that was consistent with those during the first session. However, when asked how his newly drawn circle related to the given circle, he claimed, "the size of the circles should be the same." He subsequently claimed that only "the units" differed between the two circles, reaffirming that he was not considering the circle he drew as distinct from the given circle. His drawn circle was the given circle, except with a different unit – the radius – for the measures. As Bob moved forward in the problem, he converted values from kilometers to radii and back when necessary.

Mindy provided a similar solution as Bob, except with more attention to units and dimensional analysis. She continued to stress that she trusted calculations more when she could "see" the units because "it's obvious where everything came from." We did note that Mindy used "radii" as a unit and she did not draw a distinct unit circle when solving such problems. Instead, she, like Bob, approached the unit circle as stemming from using the radius as a unit of measure. Both Bob and Mindy considered the unit circle to be a circle with *a radius of one radius length* instead of a circle with *a radius of one*.

Discussion and Implications

The difficulties that the students had at the beginning of the study suggest that their understandings of the unit circle did not support flexible reasoning about trigonometric functions or radian angle measures. Specifically, the students had numerous issues in trying to relate the unit circle and trigonometric functions to circles with a radius measure other than "one." We determined through the course of the study that the students' attempts to relate the unit circle to the given circles were not rooted in reasoning about the radius as a unit of measure, but rather the unit circle existed as a circle distinct from the given circle. Furthermore, their methods for unit conversions relied on unit-cancellation (or dimensional analysis) and did not provide a foundation for conceptualizing a circle's radius as a unit of measure or using the unit circle in flexible ways.

Our findings support Reed's (2006) observation that unit-cancellation can mask important mathematical ideas. Bob's and Mindy's actions suggest that unit-cancellation masked the quantitative meaning of calculations and provided them little support in solving unit-conversion problems that did not lend themselves to unit-cancellation. Additionally, unit-cancellation significantly restricted Mindy's reasoning, which supports Thompson's (1994) claim that, "We should condemn dimensional analysis, at least when proposed as 'arithmetic of units,' and hope that it is banned from mathematics education. Its aim is to help students "get more answers," and it amounts to a formalistic substitute for comprehension" (p. 226). Our findings also highlight the importance of magnitude-measure reasoning not only for trigonometry, but also for unit conversions and measurement.

Mindy and Bob found it unnatural to reason about the magnitude of a unit and the measure made in that unit as separate, but related objects. By prompting the students to consider units of different magnitudes and how measures change for variations in unit magnitudes, we noticed shifts in their measurement conversion schemes and their unit circle notions. The students' actions suggested that by conceptualizing a circle's radius as a unit of measure, they were able to view any given circle as the unit circle; the students no longer approached the unit circle as an object separate of a given circle. Instead, they understood that attributes of a circle, including the circle's radius, take on measures in multiple units (including radii) all at once. Stemming from this shift in their notion of the unit circle, the students more fluently used trigonometric functions in novel circle contexts by reasoning about the input and output of these functions as representing measures in radii. Weber (2005) emphasized the importance of students coming to view the unit circle as a tool of reasoning, and the influence of students' unit conversion schemes on their notions of the unit circle provides insights into how to accomplish this goal.

We note here that unit-cancellation remained a prominent way of thinking for both Bob and Mindy. However, they acknowledged (often in the moment of problem solving) that this sort of thinking hindered their progress when compared to magnitude-measure thinking. In fact, several weeks after the teaching experiment both students stressed the value they found in magnitude-measure reasoning. This occurred when Bob and Mindy designed an instructional task for a final project in their content course. They designed a task in which students i) each get a string of different length, ii) measure different classroom objects in their string lengths, iii) compare the measures that they obtain, and iv) use these comparisons to determine relationships between their string lengths. The selection and design of their task suggests that they not only valued magnitude-measure thinking for themselves, but also found such reasoning important for their future students.

The deep-rooted nature of the students' ways of thinking (e.g., unit conversion through unit-cancellation) highlights the significant impact of pre-service teachers' schooling on their mathematical content knowledge. Previous to participating in a teacher preparation program, pre-service teachers likely encounter 12 to 15 years of mathematics courses, each of which influence their content knowledge. As our study reveals, these experiences can create obstacles that must be addressed when attempting to shape the pre-service teachers' content knowledge. In working with pre-service teachers, we have the opportunity to identify ways of thinking students develop over their years of schooling and how these ways of thinking inhibit or support their learning. Such knowledge can contribute not only to improving the preparation of future teachers at the undergraduate level, but also to the mathematics education of K-12 students.

References

Akkoc, H. (2008). Pre-service mathematics teachers' concept image of radian. *International Journal of Mathematical Education in Science and Technology, 39*(7), 857-878.

Bressoud, D. M. (2010). Historical reflections on teaching trigonometry. *Mathematics Teacher, 104*(2), 106-112.

Brown, S. A. (2005). *The trigonometric connection: Students' understanding of sine and cosine*. Ph.D. Dissertation. Illinois State University: USA.

Brown, S. A. (2006). The trigonometric connection: Students' understanding of sine and cosine. In J. Novotná, H. Moraová, M. Krátká & N. Stehlíková (Eds.), *Proceedings of the 30th conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 228). Prague: PME.

Fi, C. (2006). Preservice secondary school mathematics teachers' knowledge of trigonometry: Cofunctions. In S. Alatorre, J. L. Cortina, M. Sáiz & A. Méndez (Eds.), *Proceedings of the 28th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Mérida, México: Universidad Pedagógica Nacional.

Glasersfeld, E. v. (1995). *Radical constructivism: A way of knowing and learning*. Washington, D.C.: Falmer Press.

Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A proceptual view of elementary arithmetic. *Journal for Research in Mathematics Education, 26*(2), 114-141.

Moore, K. C. (2010). *The role of quantitative and covariational reasoning in developing precalculus students' images of central concepts of trigonometry*. Ph.D. Dissertation. Arizona State University: USA.

Moore, K. C. (submitted). Making sense by measuring arcs: A teaching experiment in angle measure. *Educational Studies in Mathematics*.

Moore, K. C., & Carlson, M. P. (2012). Students' images of problem contexts when solving applied problems. *The Journal of Mathematical Behavior, 31*(1), 48-59.

Reed, S. K. (2006). Does unit analysis help students construct equations? *Cognition and Instruction, 24*(3), 341-366.

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.

Steffe, L. P. (1991a). *Composite units and their schemes*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Steffe, L. P. (1991b). Operations that generate quantity. *Journal of Learning and Individual Differences, 3*(1), 61-82.
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* (pp. 267-307). Hillside, NJ: Erlbaum.
Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks: Sage Publications.
Thompson, P. W. (1989). *A cognitive model of quantity-based algebraic reasoning*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
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 Glaserfeld* (pp. 412-448). London: Falmer Press.
Thompson, P. W. (2008). Conceptual analysis of mathematical ideas: Some spadework at the foundations of mathematics education. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sépulveda (Eds.), *Proceedings of the Annual Meeting of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 45-64). Morélia, Mexico: PME.
Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In S. Chamberlin, L. L. Hatfield & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education: Papers from a planning conference for WISDOM^e* (pp. 33-57). Laramie, WY: University of Wyoming.
Thompson, P. W., Carlson, M. P., & Silverman, J. (2007). The design of tasks in support of teachers' development of coherent mathematical meanings. *Journal of Mathematics Teacher Education, 10*, 415-432.
Topçu, T., Kertil, M., Akkoç, H., Yilmaz, K., & Önder, O. (2006). Pre-service and in-service mathematics teachers' concept images of radian. In J. Novotná, H. Moraová, M. Krátká & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 5, pp. 281-288). Prague: PME.
Weber, K. (2005). Students' understanding of trigonometric functions. *Mathematics Education Research Journal, 17*(3), 91-112.
Wildi, T. (1991). *Units and conversion charts: A handbook for engineers and scientists*. New York: IEEE Press.

SOCIOMATHEMATICAL NORMS AND MATHEMATICAL SOPHISTICATION: A QUALITATIVE CASE STUDY OF AN INQUIRY-BASED MATHEMATICS COURSE FOR PRESERVICE ELEMENTARY TEACHERS

Jennifer E. Szydlik
University of Wisconsin Oshkosh
Carol E. Seaman
University of North Carolina Greensboro

We document the evolving meanings that preservice elementary teachers ascribed to the sociomathematical norms of a mathematics class designed to foster mathematical sophistication. Specifically, we explore the developing meanings students gave to: a) their instructor's request for general solutions to problems; b) classroom norms concerning problem solving behaviors; and c) their instructor's expectation for mathematical justification. Finally, we document changes in student mathematical sophistication during the course of a semester, and illuminate the reflexive relationship between their mathematical sophistication and their interpretations of these classroom sociomathematical norms.

Key words: sociomathematical norms, classroom culture, justification

It is impossible to distinguish and thus contrast the interpretation of a thing from the thing itself … because the interpretation of the thing is the thing (Mehan & Wood, 1975, p. 69).

We describe the evolution of the culture of a highly reformed mathematics classroom for preservice elementary teachers. Our methods were borrowed primarily from ethnography; in particular, it was our attempt to understand the developing meanings the *students* ascribed to the emerging classroom norms as a way of making sense of what the students were learning about mathematics. We assumed that these meanings were conceived through primarily cultural and social processes (Cobb & Bauersfeld, 1995), and thus best observed through the lens of the classroom as a culture.

> The understanding of learning and teaching mathematics … support[s] a model of participating in a culture rather than a model of transmitting knowledge. Participating in the processes of a mathematics classroom is participating in a culture of using mathematics, or better: a culture of mathematizing as a practice. The many skills, which an observer can identify and will take as the main performance of the culture, form the procedural surface only. These are the bricks for the building, but the design of the house of mathematizing is processed at another level. As it is with cultures, the core of what is learned through participation is *when* to do what and *how* to do it…. This is to say, the core effects as emerging from the participation in the culture of a mathematics classroom will appear on the metalevel mainly and are "learned" indirectly (Bauersfeld, 1993, p. 4).

We primarily focused on sociomathematical norms of the classroom culture rather than on social norms. Yackel & Cobb (1996) defined a *sociomathematical norm* as a norm that is specific to the participants' mathematical activity. For example, while it is a social norm that a student should share an idea if it is different from that which has been previously shared, what counts as *mathematically different* in a classroom is a sociomathematical norm. Likewise, what counts as a *sophisticated solution*, a complete *mathematical explanation* or a convincing *justification* are sociomathematical norms (Yackel & Cobb, 1996). These norms tell the participants of the classroom culture *when* and *how* they should participate.

Our theoretical perspective was that of symbolic interactionism because it is compatible with constructivist learning theory, and because it embraces both social processes and sense-making processes of individuals without giving either supremacy. Blumer (1969), following ideas of John Dewey and others, advanced this view as follows:

> [Symbolic Interactionism] does not regard meaning as emanating from the intrinsic makeup of the thing that has meaning, nor does it see meaning as arising through a coalescence of psychological elements in the person. Instead, it sees meaning as arising in the process of interaction between people. The meaning of a thing for a person grows out of the ways in which other persons act toward the person with regard to the thing. Their actions operate to define the thing for the person. Thus, symbolic interactionism sees meaning as social products, as creations that are formed in and through the defining activities of people as they interact. (p. 4, 5)

Thus, it makes sense that the study of an individual's understanding of mathematics is informed by study of the culture that helped to define mathematical meanings for that individual. In other words, a participant's taken-as-shared meanings of the sociomathematical norms of a class contribute to her understanding of what mathematics is and how it is done in that class. A student who knows that her solution is mathematically different has made a meaningful mathematical distinction; a student who uses normative strategies understands what it is to do mathematics; and a student who creates a convincing argument understands what it means to justify mathematics. These distinctions are part of what Bauersfeld called the "design of the house of mathematizing." This is the level at which much learning takes place. Indeed, in a study of four elementary school classrooms, Kazemi & Stipek (2001) argued that differences in classroom sociomathematical norms accounted for differences in student performance on even traditional measures of mathematical understanding.

There is strong consensus among mathematics educators regarding broad norms that support learning, and these are articulated in the standards documents (National Council of Teachers of Mathematics, 1989, 1991, 2000). For example, more than a decade ago the *Professional Standards for Teaching Mathematics* (NCTM, 1991) called for mathematics teachers to focus on logic and mathematical evidence for verification; mathematical reasoning as opposed to memorization; on conjecturing, inventing, and problem solving; and on connections among mathematical ideas.

During the past fifteen years, the role of the teacher in bringing forth a desired culture has been described variously as negotiating meaning for language about doing mathematics (Lampert, 1990); framing paradigm cases and initiating whole class discussions of student obligations and expectations with respect to those cases (Yackel, Cobb, & Wood, 1991); offering challenge and surprise; and polishing students' verbal production (Bauersfeld, 1993). More recently Cobb, Boufi, McClain, & Whitenack (1997) described the teacher's role as "…giving commentary from the perspective of one who could judge which aspects of the children's activity might be mathematically significant" (p. 262); initiating "…shifts in the discourse such that what was previously done in action can become an explicit topic of conversation" (p. 269); and developing symbolic records of participant contributions to the discourse. Cobb et al. used the term *reflective discourse* to describe dialog that directs students to consider what was previously done in action as an object. For example, the shift could be one from generating a table of data to reflecting on the table as an object that might help students see a pattern or make an argument. A discussion of the complexity of the teacher's role is rendered beautifully in Lampert's *Teaching Problems and the Problems of Teaching* (2001) in which she makes explicit much of that which is learned indirectly. Indeed, recent studies have suggested that the teacher's role is no less than to serve as a representative of the mathematical community (Yackel & Cobb, 1996). Bauersfeld (1995) asserts that teachers must be "…exemplary, living models of the culture wanted, with transparent modi of thinking, reflecting, and self-controlling" (p. 158).

In a previous paper (Seaman & Szydlik, 2007), we proposed a partial list of sociomathematical norms of a "culture wanted," and we described an individual who embodies these as *mathematically sophisticated*. We stressed that mathematical sophistication did not imply an understanding of any specific definition, mathematical object, or procedure. Rather, it meant possessing the mathematical community's *taken-as-shared meanings for mathematical behaviors* that allow one to construct mathematics for oneself. (We note that we used the term "sophisticated" to describe an *individual* whereas Yackel & Cobb (1996) have used it to refer to the quality of a mathematical *strategy* or *solution*.)

For the purpose of discussing the mathematical sophistication of the students in this study we highlight several behaviors and values of the mathematical community. We contend that mathematicians value most highly an understanding of patterns and relationships. Steen (1990) writes, "Seeing and revealing hidden patterns is what mathematicians do best" (p. 1). Poincaré claimed, "Mathematicians do not study object, but relations among objects…" (Gallian, 1998, p. 115). Mathematicians study patterns and relationships by making and testing conjectures; by creating powerful mental, physical and symbolic models for objects, operations, and processes; and by making deductive arguments for, or creating counterexamples to, generalizations. Precise mathematical definitions of objects and relationships, and precise language and notations are also highly valued by the community of mathematicians.

The role of the instructor for the mathematics classroom in this study was to make a deliberate attempt to embody and make transparent the behaviors and values listed above. In this paper, we provide evidence that the instructor was successful in her attempt to do this; however, our primary goal is to make sense of how the students developed in terms of their own mathematical sophistication as they participated in and influenced the classroom culture. Specifically, we explore the developing meanings that students gave to the emerging sociomathematical norms of the classroom, and we compare their participation in and their verbal interpretations of the classroom culture with their mathematical sophistication, as observed through their evolving abilities to solve problems and justify solutions.

Methodology

In order to represent a variety of viewpoints and interpretations of classroom culture, the research team consisted of a mathematics education researcher who is also a mathematician (Seaman), the classroom instructor who is both a mathematician and a mathematics educator (Szydlik), and a senior-level undergraduate mathematics student. The observed class was the first of a sequence of mathematics content courses for prospective elementary teachers; there were 32 students in the class. The course content focused on number theory and arithmetic processes involving natural numbers, integers, and rational numbers. The content was constructed through daily student work on, and discussion of, carefully designed problems and activities, and study of videotapes of elementary school mathematics classrooms. The instructor was experienced in establishing the desired culture; six videotapes were made of her teaching the class, for the purposes of measuring the extent to which the classroom would be considered "reformed" by the mathematics education community, and to help the team document and make sense of classroom culture. The tapes received scores indicating highly reformed teaching practice (typically in the 80's out of 100 points) on the Reformed Teaching Observation Protocol (RTOP) (Piburn & Sawada, 2000) from both daily classroom observers and from RTOP collaborative team members. For a detailed account of the instructor's practice in this course, see Szydlik, Szydlik & Benson (2003).

The research team was present at each hour-long meeting of the course and kept daily field notes. Semi-formal interviews were conducted with six student informants four times during the fourteen-week course, and informal interviews were conducted throughout the term. The research team met weekly to discuss interpretations of classroom mathematical events and to design interview protocols. Student written work was collected throughout the term, and primary informants were videotaped solving problems aloud at both the start and the end of the semester. The informants were chosen initially based on their willingness both to participate in the study, and to share their thinking about doing mathematics. One of the primary informants needed to withdraw from the course after the first week of the term. A second informant was later replaced by the team by a student more eager to share her interpretations of mathematical work. Four informants (Lisa, Beth, John, and Andy) participated in the study for the entire semester, and we now focus our discussion on the evolving meanings they gave to the mathematical work of the class.

Results

At the start of the course, Lisa, Beth, John, and Andy were typical of elementary education majors at our comprehensive Midwest institution. They described their previous experience with mathematics as primarily memorization of procedures presented by their teachers; they exhibited weak content knowledge of arithmetic, number theory and number systems; they were profoundly mathematically unsophisticated; they were hard working and serious about their learning; and they were attentive and willing to participate in all the class activities.

In the first interview, conducted after approximately two days in the course, Beth, John, and Lisa observed that the social norms of this class were different from that of past mathematics classes. John's statement about his experience in the class was typical. "I thought we would be mostly sitting there and listening to her lecture, and not, you know, so much participation…. [Now] I think you are going to learn more things about why you're doing what you're doing; and it will make you think, not just in math, but even in the real world." Andy, however, did not recognize yet that anything novel was occurring in the course. "The teacher [in high school] would show us how to do problems and that, and we would learn how to do them, and write them down, *like we are doing now* [our emphasis]."

By the time of the second interview (week four) all four informants emphatically asserted that they had never experienced a mathematics course like this. All four informants now attributed the uniqueness of the class to sociomathematical norms rather than to social norms. In other words, while students were typically comfortable with the social norms of working on problems in small groups, sharing their mathematical work, and entertaining the mathematical ideas of others almost from the first day, it took them several weeks to recognize that something different was expected of them mathematically; they were struggling to give meaning to the mathematical expectations of the instructor.

In the following discussion we focus on the evolving meanings that these students ascribed to three specific classroom sociomathematical norms: a) the expectation that they find a *general solution* to problems; b) the expectation for *problem solving behaviors*; and c) the expectation that they *justify solutions mathematically*. Throughout the discussion, we refer to five specific problems (Figures 1 and 2), their solutions, and samples of justifications the instructor found acceptable. These problems are typical of the types of problems explored daily in the class.

Figure 1. Classroom Problems.

	Problem Statement	**Problem Solution**	**Sample of Acceptable Justification**
Poison Problem	The game of Poison is played by two teams, who take turns removing 1 or 2 counters from an initial set of n counters. The team that takes the last (POISONED) counter loses. Devise a winning strategy for any number of initial counters.	If n is a multiple of 3 plus 1, choose to go second and then take away 1 counter when your opponent takes 2, and 2 when your opponent takes 1. This will force your opponent to take the last counter. If n is not a multiple of 3 plus 1, choose to go first and take either 1 or 2 counters, whichever number will leave your opponent facing a multiple of 3 plus 1 counters.	Your opponent will lose whenever you force him to face a multiple of 3 plus 1 counters. This is because, by your taking 1 counter when he takes 2 and 2 when he takes 1, you are removing a total of 3 counters in each round. Since $n = 3m + 1$, after m rounds, your opponent will face the last poisoned counter.
Pizza Cuts Problem	Suppose you have a pizza and you get to make n straight cuts across the pizza anywhere you want. What is the maximum number of pieces of pizza you can make?	The maximum number of pieces of pizza you can make from n straight cuts is $k_n = n + k_{n-1}$, which can also be calculated by the expression $\frac{2+n+n^2}{2}$.	To maximize the number of pieces, each new cut should divide as many as possible of the old pieces in two. This will happen when each new cut intersects each of the previous cuts. The n^{th} cut will then divide n old sections in two and leave the remaining sections undivided. If we let k_n be the number of pieces at the n^{th} cut, then $k_n = 2n + k_{n-1} - n = n + k_{n-1}$. This formula generates the sequence 2, 4, 7, 11, 16, 22, 29, 37,…, the n^{th} term of which can also be expressed as $\frac{2+n+n^2}{2}$.
Number of Factors Problem	Find a way to compute the number of factors for any natural number (>1) from its prime factorization.	If $n = p^a q^b \ldots r^c$, where $p, q, \ldots, r$ are distinct primes, then the number of factors that n has is the product $(a + 1)(b + 1)\ldots(c + 1)$.	A prime number p raised to the a^{th} power has the following factors: $1= p^0, p^1, p^2, p^3, \ldots, p^a$. So p^a has $(a + 1)$ factors, q^b has $(b + 1)$ factors, …, and r^c has $(c + 1)$ factors. A factor of n could contain any of the $(a + 1)$ factors of p^a, any of the $(b + 1)$ factors of q^b, …, or any of the $(c + 1)$ factors of r^c. Thus the total number of factors of n is the product $(a + 1)(b + 1)\ldots(c + 1)$.

What did it mean to solve a problem in the class?

In the class, the expected solution to a posed problem was often a generalization that held for all natural numbers. (Students referred to problems of this nature as "number problems" or "formula problems.") The instructor described her expectation to the class regarding the nature of these solutions on the second day of class:

[Poison] (Figure 1) is a game played with some number of pennies. It could be *any* number. We're going to start with ten. But really it could be any number of pennies, so I am just going to write "*n* pennies" for a minute to stand for any number. Is there a way to play so you are guaranteed to win? I would like to know, in the end, how to play with *every* number of pennies.

Although there is evidence from the videotape and from their written work on the problem that the majority of students understood the expectation that they solve the problem for "*n* pennies," some students struggled to give meaning to the request for this type of solution. For example, on the Pizza Cuts Problem (Figure 1) a few days later, several students indicated that they did not understand that there might be a systematic way in which the number of pieces grows with the number of cuts. Consider the conversation that took place between the instructor and Lisa's group when that problem was initially posed:

Lisa: It could be any number, well not really, but you could just keep drawing lines, drawing lines, drawing lines, until you get as many as you want…

Instructor: The more cuts you make, the more pieces you can make, that's true, but if I tell you, you can only make *n* cuts, like 8, what's the most pieces you can make? So you want to answer it for some fixed *n*, it's just unknown. You see?

Lisa: Okay [Shaking her head "yes"].

Andy, too, initially struggled to give meaning to the request for general solutions. In the first weeks of the term, he interpreted a "solution that worked for any value of *n*" to be one that worked for some *specific* value of *n*, and so, for example, he addressed only the case $n = 10$ in his written work on the Poison Problem. We note that by the third week of class, Andy did give normative meaning to the expectation that he find a general solution, as evidenced by his written work on a problem of counting squares in a grid: "We were asked how we would figure out how many squares would be in a square '*N*' rows by '*N*' columns, with '*N*' being any given number." He proceeded then to express the solution as "$N^2 + (N-1)^2 + (N-2)^2 + \ldots (N-N)^2$." In his written work on the Number of Factors Problem (Figure 1) in the eleventh week of the term, he stressed, "By using this formula [his solution] you are able to find the number of factors in *any* number, small or large [his emphasis]."

We note that the class highly valued symbolic representations of, and notations for, their solutions to problems. In fact, to them, using symbolic notation made the problem presented into a *mathematical* problem. They valued most highly a closed-form, symbolic model of a solution.

What are appropriate strategies for solving a "number problem"?

In the first interview (completed at the start of the course), we asked the informants to work on the Photo Problem (Figure 2) so that we could observe the strategies they would use initially on problems like those that would be posed in the class. All four informants began by thinking aloud about the number of positions possible for each person, and after a brief time, each initially conjectured that the number of photographs of 10 people would be 100, because each person could sit in 10 places.

Lisa: I would say probably 100 times because if you've got, I mean, if you label all these people, and switch around where you had A here and A here and A here – as we said, that would be ten times. And if you switched it where it would be person B, and switched them, and switched person B ten times to be here, to be here, to be here… that would be ten times too. So if you were to go all along the row, switch them up, … I'm sure it's bigger than 100 though.

Three of the four informants (all except Andy) were unsatisfied with their initial conjecture of 100, because their intuition suggested that the number of orderings should be larger. Beth confessed that, in cases like this, she would now try to make an exhaustive list of all possible orderings.

Beth: I have a tendency to do things the long way. Like today in class, we did our problem, and I did it the total long way. I just wanted to write every number down, and it's probably what I would do in this case if I had the time. And write all the different possibilities of where – or make like a chart with x's of spots where one could be in – and then… Next I would probably put 1 where the 2 is and then put 2 where the 3 is and then 3 where the 4 is and 4 where the 5 is and so on. And then, I'd just keep doing that right down the line. And write all those down.

In the initial interview, only Lisa eventually began to collect information on smaller versions of the problem in order to test her conjecture and build a generalization for the problem.

Lisa: I've done this problem, I'm sure, in a different set-up, because I remember thinking, how many times can you do it? I'm trying to think of like a set-up, like a kind of formula. I would think to try to do a table, but, obviously doing something like this is doing a table if you do it many times, but trying to think of a formula where if I had person A, I could plug it in and find out how many times… I don't know, I can't find the formula … If it were just A and B, they could take the picture two times. And if you had person A, B, and C you could take it …[long pause]. I don't know… I am trying to figure out if each answer is increased.

While the informants had few successful strategies for solving the Photo Problem at the initial interview, class videotapes suggested that as early as the Pizza Cuts Problem (day three of class), almost all the students (including the four informants) approached finding a solution by generating data for small cases, organizing that data in a table, and then looking for patterns. They did not, however, value using the structure of the problem to explain those patterns, even though the instructor prompted them to consider why the patterns made sense and led the class through an examination of the underlying structure.

In the final interview (thirteenth week of the term), all four informants immediately used these same strategies (collecting data on small cases, organizing the information, and looking for patterns) in their videotaped work on the Circle Pattern Problem (Figure 2). Both Lisa and Beth demonstrated fairly refined ways of collecting and organizing their data. For example, Lisa first fixed n and varied m, then she fixed m and varied n. We note, however, that their ways of looking for patterns were quite naïve. First, none of the four informants used the structure of the problem to assist them. This was disappointing, because making sense of the problem in this way had become a sociomathematical norm of the class. Furthermore, in this case, the physical behavior of generating patterns on the Circle Pattern Problem is potentially meaningful -- as is an analysis of extreme cases such as (n, n) or $(n, 1)$ -- yet they did not do this either. Second, the patterns they tried to see were only those involving "even and odd" and additive or subtractive processes. They did not look for multiplicative or other relationships. We note that it was common for students in class to not "see" perfect squares, and when faced with the pattern 1, 4, 9, and 16 many saw only subsequent odd differences. This lack of number sense among students sometimes affected their abilities to solve problems using normative strategies.

Figure 2. Interview Problems.

	Problem Statement	**Problem Solution**	**Sample of Acceptable Justification**
Photo Problem	Suppose you are asked to take a group photograph of ten people – they must stand side-by-side in one row (like a line-up!). How many different pictures could you take?	10 x 9 x 8 x 7 x 6 x 5 x 4 x 3 x 2 x 1 = 3,628,800	There are 10 ways (people) to fill the first place in line. Once someone is chosen, there are 9 possible choices for the second place in line. Then there are 8 ways to fill the third spot, 7 ways for the fourth, etc., until you have just 1 person left to fill the last place in line. Since any one of the 10 choices could go with any one of the 9 choices, etc., you multiply the number of choices for each place in line together to get the total number of possible arrangements.
Circle Pattern Problem	If you have the (n,m) circle pattern [n points on a circle, connected at every m^{th} point], what conditions on n and/or m guarantee you will hit all n points?	m and n must have no factors in common, that is gcd $(m,n) = 1$	If n and m have a common factor, say k, then as you move around the circle connecting every m points, you will end up back at the starting point after $m \div k$ trips around the circle, hitting $n \div k$ points in all. So, if you want to hit all n points, then $n \div k$ must equal n, and thus $k = 1$.

What does it mean to justify your solution?

Students struggled throughout the term to give meaning to the instructor's request that they justify "*why their solution made sense mathematically*." This request was made during every class discussion and was a required section (worth almost 1/3 of the points) of each written problem report. The instructor expected that a mathematical justification be either an exhaustive one or a (typically informal) deductive argument based on mathematical structure or relevant definitions; this expectation did not change during the course of the term. What did evolve were the ways in which the students understood the expectation.

The instructor first discussed what was expected of a justification in class in the context of the Poison Problem. While the instructor was describing the expectation for the justification section for *written* work, the expectation for *verbal* arguments in class was not different. "In Section 4 [of the written reports] you address things like: *Why* do you do that to win at Poison? *Why* does that *make sense*? *Why* does that work? How do you know you are going to win *for sure*? That [Section 4] is where you make your argument."

The following justification for the Poison Problem was discussed in class with all four informants present and apparently attentive:

Class Conjecture: [written on the board] With 10, go second and then always take the opposite of your opponent.

Instructor: Why is that?

John: Because it works.

Instructor: Well, why does it work? *Why* does it work? [pause]

Daniel: Anything, like 1, 4, 7, 10, 13 … they'd all work like that because you're taking 3 away from it every time and basically you lead [your opponent] down to 1.

Instructor: All right. Wait a minute. Let's draw a picture of that. Okay. So let's draw a picture in the case of … 13. Daniel says 13 will work like that. So you're taking 3 away every time and leading down to 1 … so I'm picturing this is what you're thinking:

XXX   XXX   XXX   XXX   X [written on the board by the instructor]

3's in a turn, so you're saying 3, 6, 9, 12, and 13 pennies are right here. Okay Daniel, now explain from this picture what you're talking about.

Daniel: That no matter what they choose, if they choose 2 first, you choose 1 to make it 3 that you take away from the table [indicates first group of 3] If they choose 1, you choose 2…

To convey initial meanings students gave to mathematical justification, we consider what the four informants wrote in their justification sections of their written work on the Poison Problem. Since an acceptable justification was generated in class, we speculate that the differences in justifications given by individual students are due in large part to what each student initially gave to the meaning of "justification" itself.

In her written work, Lisa was the only one of the four informants to express the essence of the justification created during the class discussion:

When your starting number of pennies is a multiple of three plus one, you want your opponent to go first. This is because, in this scenario, you want to be the one to eliminate each row of three, always leaving the poison penny last for your opponent to pick up. Hence, you are always choosing the opposite of what your opponent decides, therefore, eliminating each row in each turn. If your starting amount of pennies is not a multiple of three plus one, you are then going to choose first. Therefore, you can eliminate the amount of pennies that leaves your opponent with a multiple of three plus one.

Lisa was also the only informant to make use of the visual model developed in class, indicating that she was able to make sense of its connection to the problem and of its relevance in justifying the generality of the solution. While others observed the visual model and dutifully copied it in their notes, they did not recognize that it illustrated the essential elements of a justification. We assert that it was not simply that the others did not understand the justification (although, for some this was certainly the case as well); they did not know that what they were hearing was important, and so they did not attend to it. Instead of providing a mathematical justification for the Poison Problem, almost all of the students in the class did one of three things:

1) They appealed to a pattern the class had found or to a process they used to work on the problem. For example, Andy wrote this as his justification section: "By trial and error, you will be able to find the same way our group found out. The pattern you find when you start with one penny up to ten, is very useful in solving this problem. By using this method, this is the only way to solve the problem."
2) They simply restated their solution, as in the case of Beth. Her justification section read: "No matter how many pennies there are to pick from at the beginning, you will win if you continue to pick the opposite of your [opponent]."
3) Some justified why the problem itself was a valuable experience. John's justification section read: "This activity was one that got the brain ticking and made us realize that we can't give up; there is a solution. It also came to me that this just goes to show how much math is involved in our lives…. It also kept us looking for new ideas and conclusions."

The activity of exploring why mathematical relationships made sense was meaningless for nearly all the students in the class at the beginning of the semester. While they did appreciate that teachers of mathematics must be able to explain "why," they did not comprehend that this "why" was dependent upon the mathematical structure underlying the problem. In addition, they began the semester with no awareness that making a mathematical argument is a necessary component of doing mathematics. Instead they looked to an external authority to validate their solutions, asking the instructor, "Is this right?"

By the fourth week of the course, all four informants had realized that the instructor expected something different as a mathematical justification. Their struggle to give meaning to that expectation is evidenced in their second interview responses.

Interviewer: How do you know when you are done with a problem?

Lisa: (Sigh) I never know when I'm done. Four weeks of class and I don't think I'm done with any of the problems… I probably [am done] when I know what I'm talking about and I can be able to explain it… I don't think finding a formula is the end of the problem, necessarily. And I think that's what I'm learning from this class too… there's always a way to check it and stuff like that…seeing if its consistent…

John: That's the one thing I have problems with in class. I never know if we're done or not…I guess if I got an answer that works for everything that I thought it would for work for, then I'm done… I try to, whatever equation I came up with that solved it or whatever, I try to put the number in there. If it goes through itself, then you say it works… I never really know if there is another answer or not.

Beth: When the whole class agrees (laughs)… no, if you're doing like a number problem, you try to make sure it works with other numbers if there's a pattern. I'll come up with an idea, conjecture, but I won't have an explanation as to why … I think that when we have an answer we are done, and then when she tries to ask for more, I get kind of confused.

Andy: When I got the answer, I plug it in the equation to see if I got it right or not. If I think I got it right then I go ask someone else if they got the same answer.

Interviewer: For your write-ups I noticed there was a justification section. What is [the instructor] looking for there?

Lisa: I think she really wants us to get the teacher aspect of the problem, how you can explain it, showing examples of making a problem … more easily able to understand.

John: I guess she's trying to look for why we got what we got, and why this is so, and why we can get from here to there… try to take a simple step so you can explain it to people so it can be recalled, so if somebody ever wants to know, it will be a little easier for them to get there.

Beth: Like in-depth reasoning of how you got … no, not the process we went through like, well yeah. And then she wants to know why that answer is correct. [On the last write-up], I worked with John and he's doing this study too, and we were really bad at it [the justification]. We came up with our answer and we wrote down why we think it's the right answer and why we think there can't be any better answers. That's about all we could do. I don't know. There's no further way we could have explained it I don't think.

Andy: Just complete answers of all your thoughts and everything, and you gotta write exactly everything you wrote for notes.

Based on these responses, we speculate that the informants generally interpreted a mathematical justification either as checking whether a formula worked using relevant numbers, or as explaining a problem in the manner of a teacher (in order to make it simpler for others to understand). We note that Andy continued to view the course primarily from the perspective of one who memorized mathematical procedures. Although he received failing grades on all the justification sections on his written work to this point, his beliefs about what it meant to do mathematics still did not allow him to view justification as problematic.

By the eleventh week of the course, while they were not particularly successful at creating mathematical justifications, the four informants (and the rest of the class) did give normative meanings to the expectation. In particular, they recognized that a justification must focus on the reasons that the mathematics made sense. (Recall that this was not understood initially by three of the four informants.) Consider their justifications for their written solutions to the Number of Factors Problem (Figure 1) explored in the eleventh week of the term. Again it was Lisa who, after attending an office hour with the instructor, made the strongest argument. She wrote the essence of a general argument as to why the number of factors of a number of the form $p^a q^b$ is $(a + 1)(b + 1)$ when p and q are distinct primes. She did this by attempting to systematically list the factors for the case of $2^3 3^3$ in the form of a table with (3 + 1) rows and (3 + 1) columns. She then argued, "In completing the array model, we see that all the possible combinations of factors are present when performing multiplication." Lisa began the course with a sense of what made a mathematical justification, and so was typically successful in recognizing the essential elements and creating them later in her written work. She was also able to identify when she did not have an acceptable justification and often sought help.

Both John and Andy, in their written work on the Number of Factors Problem, justified that their solutions worked for the case of a single, specific prime to a power. For example, John showed that 3^3 had four factors by listing them: 3^0, 3^1, 3^2, and 3^3. He then wrote, "So when looking for factors from the prime product you would add one to the exponent…" John did not explain why it made sense to multiply in the case of a product of primes to powers. Andy did recognize that this justification was important, and he attempted to provide an argument as follows: "The reason you multiplied was simple. By using the long method from above for finding all the factors of 144 [dividing 144 by successive natural numbers until all factors were generated], it came out to 15 [factors]. If you *add* $(p_1 + 1) + (p_2 +1) = ?$ [his own notation], it would come out to be 8, and you know that isn't right… so we *multiplied* and it came out to the same number of factors as the long way." Beth wrote an inductive argument, claiming her correct solution made sense because "I experimented with several different prime numbers and the solution made sense with all of them…" While Beth's justification was not acceptable to the instructor, it still suggested that she had come to see that a justification is based on making sense of mathematical work.

Furthermore, at the end of the course, the informants indicated that they *valued* mathematical justifications, both for themselves and for their future students. Consider their descriptions of "the most important things they learned in the class" or "the most important things for kids to learn about mathematics" given in their final interviews at the end of the semester.

Beth: I think the most important thing is for them to understand why things are the way they are. Like, in this class, like when we have to explain things and explain why… and why if you do it a different way, why you still get that answer.

Andy: That problems can get more complicated, that there is more to it than meets the eye.

John: The thing is [children] need to be able to do it on their own and see it for themselves, like not just to see how they got the answer, but see why they're getting the answer.

Lisa: Just that knowing how to … figure out the answer is not all that is involved, and there's more um, learning kind of theories and the reasoning why behind things, like behind why a problem works.

The informants at the end of the semester showed both that they had normative meanings for the activity of mathematical justification and that they valued the activity.

Conclusions

This study demonstrated that participation in an inquiry-based mathematics classroom culture can increase mathematical sophistication in preservice elementary teachers. Specifically, we contend that as our informants came to make sense of the classroom sociomathematical norms regarding solutions to posed problems, useful problem-solving strategies, and mathematical justification, they also came to an improved ability to do these activities, they came to value these activities, and they came to see them as essential aspects of doing mathematics.

Our informants began the course unaware that number patterns could be explained by giving structure to the mathematical problem that generated the pattern; they left the course valuing this activity and able to participate in and give normative meaning to the classroom discussions of mathematical structure. They began the course without language or meaning for the terms "conjecture," "counterexample," or "proof," and in some cases without meaning for the activity of making a generalization. For example, several students struggled to understand what it meant to solve a problem "for any value of *n*." They left the course with normative understandings for creating generalizations, and highly valuing symbolic (formulaic) representations of generalized relationships. They demonstrated that they had given meaning to the activity of making conjectures and testing them for counterexamples.

However, while they came to value an understanding of why relationships made sense mathematically (both for themselves and for their future students) through their participation in the course culture, and while they came to see deductive reasoning as an essential part of doing mathematics, they typically were not successful in doing these activities even at the end of the semester. In particular, none of the four informants used physical structure to give him or her insight into solving a problem posed in the final interview. Furthermore, the informants' abilities to see relationships and patterns in data were limited by their lack of number sense; they typically looked only for arithmetic relationships and not for multiplicative or other types of patterns in data.

As Yackel (2001) observed in her study of an inquiry-based, university-level class, we observed that the social norms of this class (expectations that students work in groups to solve problems, share their ideas both with their groups and in larger discussions, and respond to the mathematical work of others) developed quickly and were established as early as the second day of the term. Within three weeks, all the students appeared to give normative meanings to the instructor's request for a *general solution* and to her request that they *work on* (use problem solving strategies for) number problems in the class. However, the students struggled throughout most of the semester to give meaning to the instructor's request that they *justify their solutions mathematically*; most students did not negotiate normative meaning for this expectation until almost eleven weeks into the semester.

Simon & Blume (1996) have suggested the following relationship between mathematical understanding and validation: "The hearing of a logical (from the researcher's perspective) argument, which complies with the established classroom norms for mathematical justification, does not necessarily bring the other community members to the understandings of the person presenting the argument. Rather the community members tend to be limited in their sense-making with respect to the argument, in their understandings of the concepts involved" (p. 29). Based on our work, we contend that one of the "concepts involved" in justifying is the concept of *justification* itself. Students in this study were able to recognize and attempt to make sense of mathematical justifications only insofar as they could give meaning to their instructor's expectation for it. In fact, the data suggest that giving normative meaning to generalizing, doing mathematics, and justifying is prerequisite to success at each of them, and that as students came to make sense of these concepts, they improved in their abilities to do them and they began to see them as valuable.

We do not claim that understanding the concept of generalizing, doing mathematics or justifying is sufficient for success. Even though our observations of classroom discourse and written work indicated that the students began to give normative meaning to the instructor's expectation that they justify their solutions, they did not make large gains in their abilities to do so. In fact, even at the end of the semester, student justifications of their mathematical work rarely satisfied the instructor as a representative of the mathematical community. However, we assert that giving normative meanings to constructs such as finding general solutions, doing mathematics, and justifying mathematical work is a necessary condition for success at these activities, and we advocate that normative meaning for these constructs must be, and can be, actively fostered.

References

Bauersfeld, H. (1993). Teachers' pre and in-service education for mathematics teaching. *Seminaire sur la Representation*, 78. Canada: CIRADE, Université du Quebec à Montreal.

Bauersfeld, H. (1995). The structuring of structures: Development and function of mathematizing as a social practice. In, P. Steffe & J. Gale (Eds.), *Constructivism in Education*, (pp. 137-158) Hillsdale, NJ: Lawrence Erlbaum Associates.

Blumer, H. (1969). *Symbolic interactionism.* Engelwood Cliffs, NJ: Prentice-Hall.

Cobb, P., & Bauersfeld, H. (Eds.). (1995). *Emergence of mathematical meaning: Interaction in classroom cultures.* Hillsdale, NJ: Erbaum.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education*, *28*(3), 258-277.

Gallian, J. A. (1998). *Contemporary abstract algebra* (4th ed.). New York: Houghton Mifflin.

Kazemi, E. & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *The Elementary School Journal*, *102,*(1), 59-80.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29-63.

Lampert, (2001). *Teaching problems and the problems of teaching*. New Haven & London: Yale University Press.

Mehan, H., & Wood, H. L. (1975). *The reality of ethnomethodology.* New York: Wiley.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics.* Reston, VA: Author.

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

Piburn, M. & Sawada, D. (2000). *Reformed teaching observation protocol (RTOP) reference manual* (Arizona State University, Arizona Collaborative for Excellence in the Preparation of Teachers Rep. No. IN00-32). Arlington, VA: National Science Foundation.

Seaman, C. E. & Szydlik, J. E. (2005). Mathematical sophistication among preservice elementary teachers. *Journal of Mathematics Teacher Education, 10,* 167.182.

Simon, M. A. & Blume, G. W. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior, 15*, 3-31.

Steen, L. A. (Ed.). (1990). *On the shoulders of giants: New approaches to numeracy*. Washington, D. C.: National Academy Press.

Szydlik, J. E., Szydlik, S. D., & Benson, S. R. (2003). Exploring changes in pre-service elementary teachers' mathematical beliefs. *Journal of Mathematics Teacher Education, 6*, 253-279.

Tall, D. (1992). The transition to mathematical thinking: Functions, limits, infinity, and proof. In D. Grouws (Ed.). *Handbook for Research on Mathematics Teaching and Learning* (pp. 495-511). New York: Macmillan Publishing Company.

Yackel, E. (2001). Explanation, justification, and argumentation in mathematics classrooms. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the Conference of the 25th Conference of the International Group for the Psychology of Mathematics Education* (pp. 9-24), The Netherlands: Freudenthal Institute, Utrecht University.

Yackel, E. & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 459-477.

Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22*(5), 390-408.

INVERSE, COMPOSITION, AND IDENTITY: THE CASE OF FUNCTION AND LINEAR TRANSFORMATION

Spencer Bagley
San Diego State University

Chris Rasmussen
San Diego State University

Michelle Zandieh
Arizona State University

In this report we examine linear algebra students' conceptions of inverse and invertibility. In the course of examining data from semi-structured clinical interviews with 10 undergraduate students in a linear algebra class, we noted that all the students said the result of composition of a function and its inverse is 1. We propose that this may stem from the several meanings of the word "inverse" or the influence of notation from linear algebra. In addition, we examined how students attempted to reconcile their initial incorrect predictions with their later computational results, and found that students who succeeded in this reconciliation used what we termed "do-nothing function" ideas. This analysis highlights several implications for classroom practice, including a possible method to help students develop object conceptions of function, as well as the need to pay more explicit attention to often-backgrounded notational issues.

Key words: linear algebra, function, linear transformation, process/object pairs

The concept of function is central to much of secondary and undergraduate mathematics. One important context where functions appear is the study of linear transformations in linear algebra. Attesting to the importance of these two topics, the literature contains many studies on the nature and learning of function, and on student conceptions and difficulties with linear transformation.

However, to date, little attention has been paid in the literature to the extent to which students construe similarity between function and transformation. This study examines one particular aspect which students view function and transformation as similar, and the influence that students' knowledge of transformation has on their understanding of function.

Theoretical Framework

There is a robust body of literature examining the nature of students' conception of function (e.g., Sfard, 1991, 1992; Dubinsky & McDonald, 2001; Carlson, Jacobs, Coe, Larsen, & Hsu, 2002). We will first examine the literature on the epistemological nature of the function concept, then proceed to accounts of the development of the function concept.

Many researchers (e.g., Sfard, 1991, 1992; Dubinsky, 1991; Monk, 1992; Zandieh, 2000) have discussed the dual nature of the function concept. Sfard (1991) asserts that many abstract mathematical concepts, function among them, can be understood either *operationally*, as processes, or *structurally*, as objects. The operational conception is couched in the language of "processes, algorithms and actions" (p. 4), whereas the structural conception speaks of abstract and intangible objects that are fundamentally inaccessible to the senses. These two distinct yet complementary aspects of a concept are related reflexively: every process needs objects to operate upon, and eventually processes become objects that can then be acted upon by other processes.

Dubinsky and colleagues (1991; Breidenbach, Dubinsky, Hawks, & Nichols, 1992) make a similar distinction between *process* and *object* as they relate to arbitrary mathematical objects and to functions in particular. Breidenbach et al. note that quite often, "it is necessary not only to encapsulate a process to obtain an object, but also to be able to unpack or *de-encapsulate* the object and return to the process – even to go back and forth at will" (p. 267).

This ability to quickly and fluently transition between process and object views is widely taken as a mark of mathematical sophistication.

In the framework of Sfard (1991), the development of a concept typically proceeds from operational to structural, passing through three stages called *interiorization*, *condensation*, and *reification*. First, during the *interiorization* stage, the student explicitly performs a process on objects that are already familiar; in the context of function, this is typically the phase where students compute tables of functional values by explicitly evaluating functional expressions at particular numbers. Next, in the phase of *condensation*, the student gradually increases in the ability to reason about the process as a coherent whole. In a sense, the procedure becomes a "black box" that objects can be pushed through without attention to the internal workings. Finally, and usually quite suddenly, the concept undergoes a *reification* and becomes an object in its own right, able to be operated upon by other processes.

Breidenbach et al.'s (1992) account of concept development is virtually identical to Sfard's; however, the vocabulary differs in a subtle way. An *action*, or "any repeatable physical or mental manipulation that transforms objects (e.g., numbers, geometric figures, sets) to obtain objects (p. 249), is said to be *interiorized* to become a process when it becomes thought of as a whole rather than a collection of steps; this notion of interiorization is quite similar to Sfard's notion of condensation. When the process is able to be acted upon by other actions, it is said to be *encapsulated*; this idea is consonant with Sfard's reification.

Monk (1992) draws a similar distinction between *pointwise* and *across-time* views of function. Students with a *pointwise* view of functions are able to create particular input-output pairs, and to find the output for a given input, but they are less able to reason about the behavior of the function as a whole; Sfard would likely speak of such students as being in the interiorization stage. As the students grow in sophistication, they become more comfortable speaking about the overall behavior of the function, and thus develop an *across-time* view of function. This is similar to Sfard's phase of condensation: they can reason about the function as a coherent whole.

The development of the function concept from process to object is not without its difficulties. Sfard (1992) notes that many students develop the "semantically debased conception" she refers to as *pseudostructural* (p. 75). Students exhibiting a pseudostructural conception may, for instance, regard an algebraic formula as a thing in itself divorced from any underlying meaning, or a graph as detached from its algebraic representation or the function it represents. Zandieh (2000) explains a pseudostructural conception as a gestalt; that is, "a whole without parts, a single entity without any underlying structure" (p. 108). In the language of Dubinsky, a pseudostructural conception of function is an object view that cannot be "de-encapsulated," or unpacked to get at the underlying process it came from.

Much work has recently been done examining students' understanding in the field of linear algebra in general (Dorier, Robert, Robinet, & Rogalski, 2000; Hillel, 2000; Sierpinska, 2000) and of the concept of transformation in particular (Dreyfus, Hillel, & Sierpinska, 1998; Portnoy, Grundmeier, & Graham, 2006). In addition, since linear transformations are a type of function, the substantial literature discussed above on the epistemology of the function concept applies equally to linear transformation.

Student conceptions of linear transformation are often problematic. For instance, Portnoy, Grundmeier, and Graham (2006) examined whether students conceptualized linear transformations as processes or objects. They found that in general, students' views were purely operational; they saw transformations solely as "processes that map geometric objects onto other geometric objects" (p. 201).

Another study, conducted by Dreyfus, Hillel, and Sierpinska (1998), identified a tendency for students to use the term "transformation" to refer not to a mapping between vector spaces, but to the image of a vector under such a mapping. In other words, for many students, "transformation" seemed to mean the vector T(**v**) rather than the relation between **v** and T(**v**).

The present study contributes to these bodies of literature by examining the relationship between students' conceptions of function and linear transformation. In particular, we are interested in the influence that knowledge from the one context has on the other.

To help researchers discuss such influence, Hohensee (2011) borrowed from the study of language learning the notion of *backward transfer*: how prior knowledge changes as new knowledge is built upon it. This notion extends the traditional account of transfer as the influence of prior activities upon new situations. It should be noted that backward transfer is a value-neutral term; it is equally possible for new knowledge to positively or negatively change students' prior knowledge.

Methods and Background

The subjects of this study were undergraduate students chosen from a linear algebra course at a large public university in the southwestern United States. Near the end of the term, all the students in the class completed a reflection questionnaire exploring their understanding of several properties that are commonly spoken of in both the function and linear transformation contexts.

Several days after the class's final exam, ten students volunteered to participate in semi-structured hour-long interviews examining their conception of the similarities and differences between function and linear transformation. These interviews were for the most part clinical interviews (Ginsburg, 1997). Near the end of the set of 10 interviews, however, the researchers conducting the interviews (Zandieh and Rasmussen) began to develop some conjectures regarding how students might resolve a reoccurring dilemma and hence near the end of two of the interviews transitioned from a clinical interview to that of a one on one teaching experiment (Steffe & Thompson, 2000). All interviews were videorecorded and these recordings were transcribed. In addition, students' written work was retained. These videos, transcripts, and paperwork form the data examined in this analysis.

The original aim of the study was to examine how students construe similarity between topics in linear algebra and high-school algebra, and how they generalize a concept (e.g., invertibility) from one context to another. Accordingly, the topics covered by the interview were fairly wide-ranging and included injectivity, surjectivity, and geometric interpretation of compositions. This analysis focuses on students' responses to the last few questions of the interview:

- Find the inverse of $f(x) = 3x - 9$. (The interviewers were free to remind students of the typical high-school technique, as we were less interested in students' procedural skill than their conceptual understanding.)
- Find the inverse of $\mathrm{T}(\mathbf{x}) = \mathit{101{-}2}\ \mathbf{x}$.
- Four questions about the relationship between inverse in the context of linear algebra and high-school algebra.
- A set of questions about composition, both in the high-school algebra context and the linear algebra context. (These served to prime the concept of composition.)
- What will you get when you compose $f(x)$ with its inverse that you found earlier?
    - Perform the composition. Does the result match your prediction? If not, is there some reason your result makes sense?
- What will you get when you compose T(**x**) with its inverse?
    - Perform the composition. Does the result match your prediction? If not, is there some reason your result makes sense?

To analyze the data, we employed grounded theory (Strauss & Corbin, 1994). As we began examining the data, we noted that *all ten students* predicted that the composition of a

function with its inverse would yield 1. Six students (referred to as “resolvers”) were able to resolve the discrepancy between their prediction and the correct answer they later obtained, while four (“non-resolvers”) were not. This surprising result led to the following two research questions:
(1) What reasons do students give that the composition of a function or transformation with its inverse should be 1?
(2) What differentiates the mathematical thinking of resolvers from non-resolvers?

Analysis

We present case studies of three students. The first two are Jerry, a resolver, and Nila, a non-resolver; these students have been chosen to be more or less typical of their respective categories. The third case is that of Lawson, who appeared to transition from non-resolving to resolving with appropriate intervention from the interviewers. Consistent with the methodology of one on one teaching experiments, our analysis pays particular attention to the nature of the intervention and the mental operations at play that enabled him to make this shift.

Prototypical Resolver: Jerry

Our analysis begins with Jerry, who we chose as a representative of the resolving group. At the time of this interview, Jerry was a senior in computer engineering, taking linear algebra as a major requirement.

Jerry was first asked to find the inverse of the function $f(x) = 3x - 9$. At first, he couldn’t quite remember how to do it, so the interviewers reminded him of the common algorithm, to interchange x and y and solve for y. (This was common practice across all the interviews, as the interviews were more focused on student thinking than on whether a student remembered the procedure for finding the inverse.) Even after this reminder, Jerry seemed hesitant (“I don't really want to try”), but with a little more encouragement from the interviewers, he came up with the correct answer without further difficulty.

Next, Jerry was asked to find the inverse of the transformation T(**x**) = *101−2* **x**. As he began writing down the matrix, Jerry said, “I don’t know what the **x** is there, should I just block it?” The interviewer seemed confused by this question, replying, “I guess, for now,” in a somewhat questioning tone. By “block it,” Jerry apparently meant to ignore it for the purpose of this calculation; as he finishes writing down the matrix and augmenting it with the identity, he does not write the **x** anywhere. It will be seen later in this excerpt that Jerry is able to make meaning of the **x** in equations defining linear transformations. However, it appears that despite the instructor’s best efforts, Jerry has not appropriated the practice of appending **x** to his matrices in matrix calculations. Jerry proceeded to correctly solve the problem, row-reducing the augmented matrix and reading the inverse matrix off the right-hand side.

Jerry’s initial prediction for the result of the composition of a function and its inverse was 1, because “they sort of cancel each other.” He was then asked to carry out the calculation, coming up with the correct answer, x.

Michelle: Does that surprise you?

Jerry: The whole cancelation thing doesn't surprise me, but my original thought was…

Michelle: 1? Do you think it's weird that it should be x when your initial guess was 1, or is there some reason why it makes sense for it to be x?

Jerry: Well, hm. No, whatever you put into it, that's what you're going to get out of it. This I'm thinking about with the x, whatever x you have, put into the function. And then run it again with its inverse, you're pretty much just going back to x.

Jerry's resolution came as he expressed what we have come to call *do-nothing function* (DNF) ideas, expressing the result of the composition of a function with its inverse as the function that does nothing. Reasoning with the DNF appeared to be particularly powerful in students' attempts to reconcile their incorrect prediction with their correct answer. DNF ideas were a mark of more productive reasoning; remarkably, each of the six resolvers, and none of the four non-resolvers, used the DNF to explain their result.

Here, Jerry clearly construed the function $f(f^{-1}(x)) = x$ as the DNF: "whatever you put into it, that's what you're going to get out of it." Since the function does something and the inverse undoes it, "you're pretty much just going back to x" – i.e., taken as a whole, the composition does nothing. The importance of DNF ideas in Jerry's reasoning became even more apparent when the interviewers revisited this theme at the end of the interview:

> Chris: Any final thoughts on your original prediction for f composed of f inverse to equal 1? […]
>
> Jerry: How 1 would work out? I just sort of saw it as canceling, just a bunch of canceling each other out, you end up with just 1 by it-. Uh.
>
> Chris: Is the canceling like f and f to the negative 1st, like f over f, do those cancel, is that what's canceling to give you 1? Or is it something else canceling?
>
> Jerry: […] Yeah, I see the functions canceling. But the, I don't know, now it just makes more sense that's whatever you put in there, is whatever you're getting out.

Jerry explained here that he *originally* saw the function and its inverse as canceling to yield 1. *Then*, however, he decided that x is a more reasonable answer, again because "whatever you put in there, is whatever you're getting out." Jerry thus appears able to view the result of the composition process (i.e., x) as a function in its own right, and to view this function as the function that does nothing.

Now that Jerry had thought about the composition $f(f^{-1}(x)) = x$ as being the do-nothing function, he became able to make a correct prediction about the composition of a transformation and its inverse. He explained his prediction in DNF terms:

> Michelle: If you compose T and T inverse, so similar to this but T's, T inverse, what would you predict that you'll get?
>
> Jerry: I, what's a, just **x** again.
>
> Michelle: You think you might get **x** again, how come? Just because you have **x** here, or some other reason?
>
> Jerry: We kind of did this in class. You're pretty much transforming it into something else, and the inverse really just transforming to, or transforming it back to what it originally was.

Jerry's language here echoes his prior language: "transforming it into something else, and… transforming it back to what it originally was" sounds quite similar to "put[ting x] into the function … then run it again with its inverse, you're pretty much going back to x." In both cases, Jerry appears able to reason about the result of the composition of two functions (or transformations) by pushing an arbitrary element through and observing that it will not change. He recognizes the result of composition of two functions (or transformations) as a function (or transformation) in its own right, and reasons that since it does nothing to an arbitrary element, it must be the do-nothing function. It will be seen later that this is the calculation Jerry carries out in the case of transformation. In particular, it is clear that Jerry is not simply parroting the result from the previous problem, but rather that this prediction comes from a deep understanding of the process of composition of a transformation and its inverse.

Jerry proceeded to work through the calculation in something of an unorthodox way, pushing an arbitrary vector through the calculation rather than multiplying the matrices first:

$$\begin{bmatrix}1 & 0\\ \frac{1}{2} & -\frac{1}{2}\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}x+0\\ \frac{x}{2}-\frac{y}{2}\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\ 1 & -2\end{bmatrix}\begin{bmatrix}x\\ \frac{x}{2}-\frac{y}{2}\end{bmatrix} = \begin{bmatrix}x+0\\ x-2\left(\frac{x}{2}-\frac{y}{2}\right)\end{bmatrix}$$

$$= \begin{bmatrix}x\\ \cancel{x-x}+y\end{bmatrix} = \begin{bmatrix}x\\ y\end{bmatrix}$$

Fig. 1: Jerry's calculation

Jerry was satisfied with this result; when asked if this is what he expected to get, he replied in the affirmative. It is noteworthy that his unorthodox method of calculation here appears to come from the same place as his previous hypothetical reasoning about the composition process: in both cases, he pushed an arbitrary element through and reasoned about the result of the composition by examining its effect on the element.

In both the function case and the transformation case, Jerry concluded, when you compose with the inverse, "you end up with the same input." It seems clear that the do-nothing function has provided him with a useful way to think about composition with the inverse; in particular, with the help of DNF ideas, Jerry appeared able to unpack the function or transformation object to obtain and reason with the underlying process. This ability to "de-encapsulate the object and return to the process" (Briedenbach et al., 1992) is widely recognized as a mark of a sophisticated conception of function.

Prototypical Non-Resolver: Nila

We next consider the case of Nila, who we chose as a representative of the students who were unable to resolve the discrepancy. At the time of this interview, Nila was a sophomore majoring in mathematics. Nila was first asked to find the inverse of the function $f(x) = 3x - 9$, which she did quickly, correctly and confidently, without needing a reminder of the procedure. She hesitated only slightly when writing down the formal notation $f^{1}(x)$, seeking confirmation that this is the proper notation to use.

Nila's next task was to find the inverse of the transformation $\mathrm{T}(\mathbf{x}) = 101{-}2\ \mathbf{x}$. She performed the typical calculation, augmenting the matrix with the identity and row-reducing. Just like the previous one, this calculation did not appear to cause her any significant difficulty.

Nila was next asked to predict the result of composition of a function and its inverse:

Nila: Oh, so you're saying if I put these together and then? Okay. Oh, in that case, if you take this one [points to $f(x) = 3x - 9$] and multiply it by this one [points to the inverse], it's supposed to give you 1 or is it -1? I forgot. I think it's 1, let me see.

Michelle: So you're going to multiply them?

Nila: Yeah, I'm going to multiply them. Yeah, I think it's supposed to give me 1.

This is an example of the common confusion between multiplication and composition. This confusion is seen in all of the non-resolvers and none of the resolvers. However, as Nila further examined the expression, she became unconvinced that multiplication will give her the result she wants, because as the interviewer pointed out, the result is "going to be x^2 and a bunch of ugly stuff."

To refresh Nila's memory on the topic, the interviewers proceeded through the questions on composition. They then asked her again to predict the result of composition of $f(x)$ and its inverse. She repeated her initial prediction: "I think that would work out to be 1." The interviewers had her carry out the composition; she reached the correct answer, x, but seemed startled:

Michelle: So you're surprised you got x instead of 1?

Nila: Um-m-m!

Michelle: Or is that a good thing that you got x?

Nila: [Emphatically] I have no idea.

Michelle: It is the right answer.

Nila: I don't know why I was thinking 1, but I was thinking 1.

The transcript of this portion does not do justice to Nila's emotional expressions. While making the distressed noise here represented as "Um-m-m!", she moved her hands as if pushing away the offending paper, and the tone of her voice suggested hostility, as if the problem had tricked or betrayed her. She did not even attempt to reconcile this result with her incorrect prediction. The interviewers, perhaps in response to her obvious distress, moved on to the next question without pressing her further.

Nila's next task was to predict the result of composition of a transformation with its inverse. She predicted that the result would be the identity matrix, but as time had run out and the next student was waiting, the interviewers did not have her carry out the calculation.

Transitioner: Lawson

Finally, we will examine the case of Lawson, a senior in computer science at the time of this interview. Lawson's case is particularly interesting because of the transition he appeared to make with the intervention of the interviewers from a non-resolving to a resolving position.

Lawson's interview proceeded similarly to Jerry's and Nila's. Although he had expressed earlier in the interview a confusion between the algebraic inverse (i.e., the reciprocal) and the functional inverse of a function, he applied the standard procedure of switching x and y and solving for y when asked to find the inverse of $f(x) = 3x - 9$, and found the inverse correctly. He explained that his confusion between these two concepts stemmed from the common notation used to represent both: "I just remember the inverse notation being this [a superscript -1], and I think I just automatically applied that for some reason."

Lawson was next asked to find the inverse of the transformation $T(\mathbf{x}) = \mathit{101{-}2}\ \mathbf{x}$. He made a false start by augmenting it with the zero vector rather than the identity matrix, but quickly realized and corrected his mistake. The rest of his calculation proceeded quickly and accurately.

When asked to think in general about the composition of a function or a transformation with its inverse, Lawson said, "whenever I see something like this [i.e., $f(f^{1}(x))$ or $T(T^{-1}(\mathbf{x}))$], I automatically just want to cancel them out, make them 1 or something." This provides further evidence of Lawson's conflation of algebraic and functional inverses.

Lawson reiterated this prediction for the result of the composition of $f(x)$ with its inverse, but when he carried out the calculation, he correctly obtained x. He did not appear to be as surprised or shocked as Nila, but still could not see a way to reconcile this answer with his prediction:

Michelle: Does it surprise you that you get $y = x$?

Lawson: It doesn't surprise me, I guess, I'm not really fresh in mathematics, I would say! Linear algebra doesn't really take me back to anything I learned in the past, and I haven't done any normal algebra for a long time, so.

Michelle: See, you initially predicted it would be 1, it turns out to be an x; do you have any way of thinking about why it's x instead of 1?

Lawson: [thinks] No.

He made a few stumbling attempts to make sense of his answer, but did not arrive at anything he appeared to find satisfying, so the interviewers moved on to the next question. Lawson was next asked to predict the result of composition of T with T^{-1}:

Lawson: I assume it's going to equal the identity.

Michelle: Okay, so let's see if it does?

Lawson: [writes]

Michelle: Was there a reason you assumed it was going to be the identity?

Lawson: I think originally, because I thought of reciprocals. When I tried to figure it out this way originally, I thought it was like this [writes A/A].

Once again, Lawson exhibited a conflation of the different notions of inverse; here, as evidenced by his division notation, he appeared to confuse the algebraic inverse (i.e., the reciprocal) with the inverse of a matrix.

Near the end of the interview, the interviewers shifted from clinical interviewing to conducting a one on one teaching experiment. Earlier in the interview, they had brought to Lawson's attention the mismatch between his two results (i.e., x and the identity matrix); he had tried to reconcile this discrepancy, but was unsuccessful. At the end of the interview, the interviewers circled back and pressed Lawson a little harder on this point. Chris pointed out an important difference between the way Lawson symbolized the two problems:

Chris: When you did f composed with f inverse, the input variable x was always present. When you did the T composed T inverse, what was the input for transformations wasn't present.

Lawson: [nods]

Chris: So that seems to be an important difference. So I'm wondering what in your mind is the role of the input in the T composed T inverse? And how can you think about the role of the input, the things that you input into transformations, as you're thinking about computing T, T inverse?

Lawson: In this case [functions], I have, I'm plugging this into where x was, because x is present. Whereas in this case [transformations], it's not present. I'm not sure what you mean by 'not present' necessarily because we have **x** here [underlines the **x** in T(**x**) = *101−2* **x**].

Michelle: It's there, but it's not here [points to matrix computation]. […] When you wanted to find out what this is, it didn't appear any more.

As a first attempt to help Lawson reconcile this difference, Michelle had Lawson generate an expression for T^{-1}(**x**) parallel to the one printed on the paper for T(**x**):

Michelle: Here's a question for you: Write for me here T inverse of **x** equals, now fill in the blank. [Portion omitted] Like how here we have T of **x** equals something, so we want T inverse of **x** equals something.

Lawson: Oh, okay. You could note it [writes T^{-1}(**x**) = *10½−½*; pauses]. And I would say you can put the **x** here [writes in **x** on the right], I guess.

[…]

Michelle: So what happens if I have the **x** there, does that change what's on the right side?

Lawson: [adds **x** to the end of the matrix calculation] I would assume it has an **x** there. So that's [shrugs] the identity times **x**. Which, will that come out as the $y = x$ equation?

Even after making this observation, Lawson appears unconvinced, using a questioning tone in the above portion of transcript and agreeing with Chris that his difficulties weren't resolved yet. He said he wished that he had "some kind of revelation," but clearly did not.

Chris then attempted a similar strategy, but with a subtle difference:

Chris: So what does this mean to say 1, 0, ½, -½ times the vector **x**? Well, it means 1 times the 1st component plus, etc., right?

Lawson: Um-hm.

Chris: So if you wrote that out in terms of the symbols x and y, then you would be able to say, 'Now I need to use that and get acted on by the external function.' So I'm curious, could you just push the notation that way a bit?

With this bit of impetus from the interviewers, Lawson proceeded through a calculation similar to the one performed by Jerry. This portion of transcript occurs at the end of this calculation:

Lawson: [writes] You end up with x, y. This cancels, x would cancel, and this [y] would be positive.

Chris: So going back to what you were originally computing, was T composed T inverse of the vector x, y. And this equals?

Lawson: Essentially this is the vector **x**, so essentially I did end up with, when I composed them, I ended up with **x** as in the, whatever I had here. Yeah, it is identical. [pause] That's cool! [Laughs] I'm glad I did that, that's interesting.

It should be noted that Lawson's resolution, and his sudden feeling that the process made sense, came at the very moment he expressed DNF ideas: "I ended up with **x** as in the, whatever I had here [before the calculation]." This is further evidence of the utility and power of DNF ideas in students' explanations.

Discussion

These case studies illustrate two reasons students predict that $f(f^{-1}(x))$ should be 1: first, conflation with multiplicative inverses, and second, backward transfer (Hohensee, 2011), or the influence of linear algebra on students' prior knowledge. Additionally, they illustrate the usefulness of DNF ideas in students' reconciliation processes. In this section, each of these themes will be discussed in greater depth.

Conflation with multiplicative inverses

In the foregoing analysis, we have highlighted several examples of conflation of the various concepts all called "inverse" and all symbolized with a superscript -1. In particular, students participating in this interview discussed three distinct mathematical objects all symbolized this way: the multiplicative inverse of a number (i.e., the reciprocal), the functional inverse, and the multiplicative inverse of a matrix. Nigel, for instance, said: "So say you have x, the inverse is x to the negative 1, or 1 over x." He later wrote down two other multiplicative inverses.

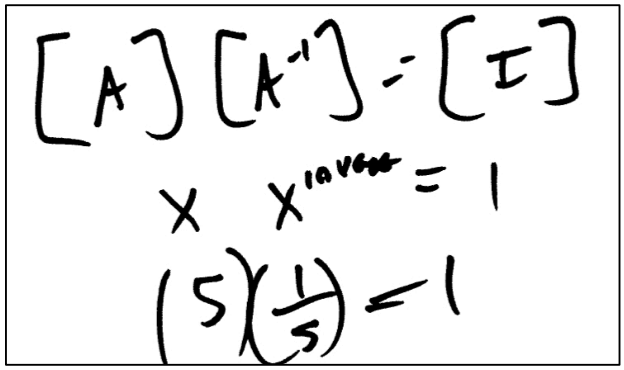


Figure 2: Nigel writes three different inverses

Additionally, students often conflate composition with multiplication; for instance, Nila demonstrated this confusion before being reminded how to compose functions. This may be due to the influence of linear algebra. The composition of two transformations is computed by multiplying matrices; this may lead students to attempt to compose functions by multiplying them. This is an instance of backward transfer, but it is not the only one.

Backward transfer

When asked to predict the result of composition of f with f^{1}, some students seemed to know the right answer but simply symbolized it incorrectly. For instance, Gabe offered the following explanation of what the result should be:

Michelle: If I do f of f inverse of x, what do you expect it to come out with?

Gabe: Input, the input that you put in there. It shouldn't modify it.

Michelle: If I haven't put in any input though, I'm just doing a calculation?

Gabe: [writes] It's just 1.

Michelle: It would be 1?

Gabe: It's not going to change what you put in there, because if you do something and then you undo it, has it really changed? It's like philosophy right there, it's going to be the same number in terms of, put in a 5, you're going to get out a 5.

Gabe appeared to know that the right answer is the do-nothing function: he explained that the composition "shouldn't modify [the input]" and that "it's not going to change what you put in there." He further illustrates this understanding with a specific example: "put in a 5, you're going to get out a 5." Accordingly, we conclude that the only reason he didn't predict x, as we may otherwise have expected him to, is because he chose the wrong notation.

Why would Gabe and other students symbolize the do-nothing function incorrectly? This may be another instance of backward transfer from the symbolism of linear algebra. It is common practice in linear algebra classrooms to omit the **x** and work directly with matrices when performing calculations on linear transformations. Thus, students are likely used to seeing the identity *matrix* alone when the identity *transformation* is under discussion. From here, it is no great leap to imagine students thinking that 1 is the identity function in the context of high-school algebra:

Gabe: If you get T, and you multiply T by its inverse, you should get the identity matrix, which is essentially 1.

Michelle: So you see those as the same?

Gabe: Yeah. 1 in matrix algebra looks like this [the identity], same thing.

To lend further support to this hypothesis, several students (including Gabe) reconciled their prediction with their result by explaining that x is the same thing as 1 times x. This is exactly how the identity transformation works; it is the identity function (analogous to 1) times the input vector (analogous to x).

DNF ideas

As mentioned earlier, all six of the students who resolved, and none of the four who did not, used DNF ideas in their explanation. Why were DNF ideas such a reliable indicator of students' ability? One plausible explanation is rooted in the dual nature of the function concept. As attested by Dubinsky (1991), the composition of functions is a rather complicated psychological problem:

> Composition is a binary operation which means that it acts on two objects to form a third. Thus, it is necessary to begin with two functions, considered as objects. The subject must "unpack" these objects, reflect on the corresponding processes, and interiorize them. Then the two processes can be coordinated to form a new process that can then be encapsulated into an object which is the function that results from the composition. This is much more complicated than simple substitution. (p. 12)

In our particular case, the function and its functional inverse must be viewed as objects acted upon by the process of composition. Algebraically, this process yields x. In order to make sense of this result, the student must unpack the function and its inverse, coordinate them by realizing that the one "undoes" the other, recognize the result of this coordination as an object (in particular, the do-nothing function), and realize that the symbol x is a reasonable way to represent this result. This requires several switches between process and object views of function. The ability to switch fluently between these views is commonly recognized as an important mark of mathematical sophistication (Breidenbach et al., 1992; Sfard, 1992). Considered in this light, it is unsurprising that DNF ideas are linked to the ability to resolve, since they entail the ability to switch views of function.

Pedagogical Implications

We find that the data discussed above tell us two stories. The first is a cautionary tale: a lack of explicit attention to notational aspects of linear algebra may foster pseudostructural conceptions of linear transformation. The second is a recommendation: helping students foster DNF ideas may help them foster sophisticated object views of function and transformation.

A cautionary tale

Educators certainly do not want students leaving an undergraduate linear algebra class saying that $f(f^{1}(x)) = 1$. How can this be avoided? One suggestion that emerges from our data is to be careful with notation. It is common practice in linear algebra classrooms to omit the **x** and write only the matrix when discussing linear transformations. Without explicit attention to this practice, students may identify the matrix with the transformation it represents. This may lead to a pseudostructural conception (Sfard, 1992) of transformation: if students conceive of a transformation as a matrix, they necessarily have an object-like view that cannot be unpacked to reveal the underlying process as applied to vectors.

Even when teachers are conscious of the possible problem, this slight abuse of notation is so common that it is difficult to eliminate entirely from their practice. In conversations with the instructor of the students we interviewed for this study, she indicated that she made a conscious effort to always write the **x** and to be clear about the difference between the matrix and the transformation it represents. Even though she was aware of these issues, she remembered several occasions where she "slipped up" and spoke of a matrix as a transformation, or omitted the **x** in board-work. The impact of such "slip-ups" can be mitigated if teachers explicitly discuss the practice of omitting **x** or speaking of the matrix as if it were the transformation, thus moving to institutionalize these as classroom mathematical practices (Stephan & Rasmussen, 2002).

As noted previously, students in this study referenced three different mathematical objects called "the inverse:" the multiplicative inverse, the functional inverse, and the (multiplicative) inverse of a matrix. To help reduce the confusion between these concepts, teachers might explicitly address the similarities and differences between them. Linear algebra is an opportune place for this discussion to occur, as it is typically students' first exposure to more general fields where the "multiplicative" inverse is analogous to, but different from, the reciprocal in the field of real numbers.

A recommendation

It appears that DNF ideas are a particularly useful way to characterize the identity function; indeed, this may even be the most useful or most "correct" way. Calling the identity function "the function that does nothing" unpacks the object to reveal the underlying process. Teachers might use these ideas to help students develop an intuitive grasp of the identity concept.

Additionally, teachers may be able to leverage DNF ideas to help students develop object conceptions of function and transformation. Consider, for instance, the interviewers' intervention that led to Lawson's resolution. They asked Lawson to write the transformation as a matrix with an **x** attached, then to push an arbitrary element through the calculation. Once he had done this computation, the interviewers encouraged him to reflect on the result, and he was suddenly able to make sense of the result.

The dramatic success of this intervention can be explained by appealing to Sfard's (1991) framework of the development of the function concept. Pushing an arbitrary element through the calculation encouraged interiorization; Lawson was attending to the specific details of the calculation. Asking him to reflect on the result encouraged condensation, since he could now ignore the fine details of the calculation and instead focus on the big-picture relationship between the beginning and ending state. Now that interiorization and condensation have occurred, reification is possible, and it appears that Lawson's sudden resolution is an instance of just that. Further research might examine whether this intervention would be as successful with other students, or in a full-class setting.

References

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics, 23*, 247–285.

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, 33* (5), 352-378.

Dorier, J.-L., Robert, A., Robinet, J., & Rogalski, M. (2000). The obstacles of formalism in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 85-124). Dordrecht: Kluwer.

Dreyfus, T., Hillel, J., & Sierpinska, A. (August, 1998). Cabri-based linear algebra: Transformations. Paper presented at the First Conference on European Research in Mathematics Education, Osnabrück, Germany.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 95-126). Dordrecht: Kluwer.

Dubinsky, E., & McDonald, M. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton, M. Artigue, U. Krichgraber, J. Hillel, M. Niss, & A. Schoenfeld (Eds.), *The teaching and learning of mathematics at university level: An ICMI Study* (pp. 273-280). Dordrecht: Kluwer.

Ginsburg, H.P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. New York: Cambridge University Press.

Hillel, J. (2000). Modes of description and the problem of representation in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 191-207). Dordrecht: Kluwer.

Hohensee, C. (2011). *Backward transfer: How mathematical understanding changes as one builds upon it* (Doctoral dissertation). Retrieved from doi:10211.10/1441

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 no. 25, pp. 175-193). Washington, DC: Mathematical Association of America.

Portnoy, N., Grundmeier, T. A., & Graham, K. J. (2006). Students' understanding of mathematical objects in the context of transformational geometry: Implications for constructing and understanding proofs. *Journal of Mathematical Behavior 25*, 196-207.

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.

Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification – The case of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (MAA Notes, no. 25, pp. 59-84). Washington, DC: MAA.

Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 209-246). Dordrecht: Kluwer.

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 (*pp. 267-307). Hillsdale, NJ: Erlbaum.

Stephan, M. & Rasmussen, C. (2002). Classroom mathematical practices in differential equations. *Journal of Mathematical Behavior, 21*(4), 459-490.

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.

Zandieh, M. J. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. *Research in Collegiate Mathematics Education, IV*, 103–127.

IMPROVING STUDENT SUCCESS IN DEVELOPMENTAL ALGEBRA AND ITS IMPACT ON SUBSEQUENT MATHEMATICS COURSES[1]

William O. Bond and John C. Mayer
University of Alabama at Birmingham

One direction taken by course reform over the past few years has been the use of computer-assisted instruction, often applied to large-enrollment service courses, and justified in part by cost-effectiveness. Elementary algebra is typically taken by undergraduate students who do not place into a credit course. The goal of such a developmental algebra course has been to enhance students' "algebra skills," for example, dealing procedurally with rational expressions. Higher-order thinking may be largely absent from such an approach. Our motivating question is "What approach maximizes the student's chance to succeed in subsequent courses?" In view of our theoretical perspective that an inquiry-based approach enhances learning, a subsidiary question is "Is it effective to blend a focus on skills development (through computer-assisted instruction)with a focus on problem-solving (through cooperative group learning)?" Results of the analysis suggest that effectiveness is a matter of what student outcomes are valued, balanced against cost-effectiveness.

Key words: developmental algebra, collaborative group work, computer-assisted instruction, quasi-experimental study, inquiry-based class meetings.

Research Question

Three studies (Mayer 2009, 2010, 2011) relevant to the current research compared treatments using quasi-experimental designs. The fundamental difference between the treatments in the two studies of a developmental algebra course (2010, 2011) was (1) incorporating one or more inquiry-based class meetings, or (2) incorporating lecture class meetings, both together with a common computer-assisted learning component. In the current research, which uses additional data gathered on the algebra student cohorts, we ask the question, "Does the treatment have a statistically significant effect on student success in the next mathematics course taken?"

Theoretical Perspective

Our research is based on the premise that active learning (Prince 2004) promotes retention of knowledge, concept development, and problem-solving (Marrongelle and Rasmussen 2008). We take computer-assisted instruction, a form of active learning, as a *ground* – the *figure* is blending with another type of active learning: inquiry-based learning (IBL) in the form of collaborative small group work and whole-group sharing. We comment here only on the figure.

In their extensive report on the IBL Mathematics Project, Laursen (et al. 2011) identifies several features of IBL "typical of their project." These features correlate well with the dimensions of the RTOP instrument for classroom observation (RTOP 2010, Sawada 2002).

---

[1] This research was supported by grants EHR-0632522 and DUE-0928665 to UAB, representing the Greater Birmingham Mathematics Partnership, from the National Science Foundation (NSF). Any opinions expressed in this article are those of the authors and do not necessarily reflect the views of NSF.

Where Laursen identifies features of the *course*, we modify this and list features of the *class meeting*:

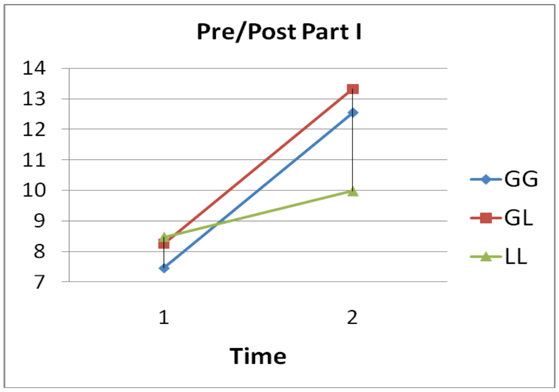
1. The main work of the class meeting is problem-solving (e.g., Savin-Baden and Major 2004; Prince and Felder 2007).
2. Class goals emphasize development of skills such as problem-solving, communication, and mathematical habits of mind (e.g., Duch, et al. 2001; Perkins and Tishman 2001).
3. Most of the class time is spent on student-centered instructional activities, such as collaborative group work (e.g., Gillies 2007; Johnson, et al. 1998; Gautreau and Novemsky 1997).
4. The instructor's main role is not lecturing, but guiding, asking questions, and giving feedback; student voices predominate in the classroom (Alrø and Skovsmose 2002).
5. Students and instructor share responsibility for learning, respectful listening, and constructive critique (e.g., Goodsell, et al. 1992; Lerman 2000; Prince 2004).

The inquiry-based treatments (described below) were designed to incorporate these features.

Prior Research and Relation to Literature

Three recent studies (Mayer et al. 2009, 2010, 2011), simultaneously compared different pedagogies over one semester. There are few such direct comparisons in the literature (examples: Doorn 2007, Gautreau 1997, Hoellwarth 2005; literature review: Hough 2010a, 2010b). Nearly all previous studies have focused on courses at the calculus level and above (Hough 2011a). The setup for the experiments was to have students sign up for a class and a time slot. The students in these fixed time slots were then randomly split into either two or three groups depending on the number of experimental teaching treatments being evaluated. This allowed each time of day to be taught with all of the different treatments to avoid a time of day effect. In Fall 2008 a quasi-experimental study was performed with finite mathematics classes at UAB. The two time slots were split into three treatment groups which were: (1) one collaborative group meeting and one computer lab meeting, (2) one lecture and one computer lab meeting, or (3) one lecture meeting with weekly quiz and one computer lab meeting. In Fall 2009 a similar experiment was performed on two classes of basic algebra but omitting the lecture/quiz treatment, and in the Fall 2010 the experiment was conducted on three classes of basic algebra with three treatments, described more fully below. The results of the quasi-experimental studies of a finite mathematic course (2009), and of an elementary algebra course (2010, 2011) showed in all cases that students in the inquiry-based treatment(s) did significantly better ($p<0.05$) comparing pre-test and post-test performance in the areas of problem identification, problem-solving, and explanation as measured by an open ended, free response pre/post test (see Figures 1 and 2). Moreover, students, regardless of treatment, performed statistically indistinguishably when compared on the basis of course test scores.

Outcomes of the first two studies by Mayer differed in gain in accuracy, pre-test to post-test: in the finite mathematics study, there was no significant difference between treatments, but in the first elementary algebra study there was a significant difference between treatments in favor of the inquiry-based treatment. In those studies, accuracy was assessed on a small set of open-ended problems. In the second elementary algebra study, the pre/post-test had both an open-ended and an objective portion. There was no significant difference among treatments in the second elementary algebra study with regard to the objective part of the pre/post-test.



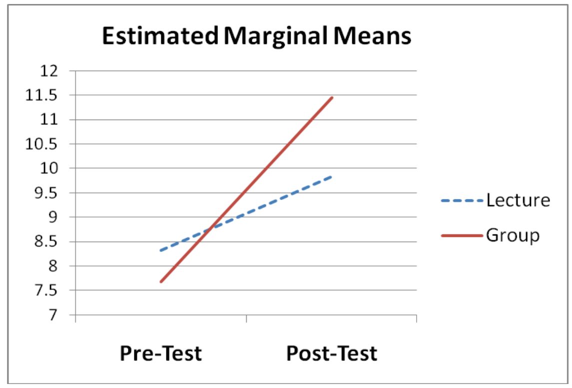


Figure 1. Open-Ended Pre/Post-Test 2010 Cohort: N=272, GG =85, GL=93, LL=94.

Figure 2. Open-Ended Pre/Post-Test 2009 Cohort: N=234, Lecture=115, Group=119.

Mayer (2011) reported that students were distinctly more satisfied with a pedagogical approach that included at least some lecture meetings (see Figure 3). Two of the three instructors received student survey scores in the lowest 10% when using the GG teaching format. It is relevant to note that these same instructors scored in the middle 40% for each of the other two treatments.

Fall 2010 Cohort: IDEA Ratings of Instruction									
Treatment	**GG**			**GL**			**LL**		
	Raw Average		Conver-ted Score	Raw Average		Conver-ted Score	Raw Average		Conver-ted Score
	Excellent Teacher	Excellent Course		Excellent Teacher	Excellent Course		Excellent Teacher	Excellent Course	
Instructor1	2.4	2.2	**25**	3.7	3.6	45	4.3	3.9	51
Instructor2	4.2	4.0	52	4.3	4.0	52	4.3	4.0	50
Instructor3	2.5	2.6	**30**	4.8	4.1	56	4.3	4.0	49

Figure 3. IDEA Survey: converted scores in the range 45-55 place instructor/course in the middle 40% of all IDEA mathematics student ratings; scores 37 or lower, in the lowest 10%.

Research Methodology

The methodology in (Mayer 2010, 2011) was quasi-experimental in that it sought to remove from consideration as many confounding factors as possible, to assign treatment on as random a basis as possible (constrained only by students being able to choose the time slot in which they take the course), and then to compare results for the same cohort of students.

All students involved in the courses had identical computer-assisted instruction provided in a mathematics learning laboratory. Each class, regardless of treatment, met once a week in the computer lab with their instructor. Homework, quizzes, and tests were all computer-based.

This methodology was described completely in (Mayer 2010, 2011). For completeness herein, we briefly describe the experimental set-up. Students registered for one of three time periods in the Fall 2010 semester schedule for two 50-minute class meetings and one 50-minute required lab meeting. Students in each time slot were randomly assigned to one of the three treatments for the semester:

(1) [GG] two sessions weekly of inquiry-based collaborative group work (random, weekly changing, groups of four) without prior instruction, on problems intended to motivate the topics to be covered in computer-assisted instruction;
(2) [LL] two sessions weekly of traditional summary lecture with teacher-presented examples on the topics to be covered in computer-assisted instruction, and
(3) [GL] a blend of treatments (1) and (2), with one weekly meeting traditional lecture, and one weekly meeting inquiry-based group work.

Students registered for one of four time periods in the Fall 2009 semester schedule for one 50-minute class meeting and one 50-minute required lab meeting. Students in each time slot were randomly assigned to one of the two treatments for the semester, similar to (1) designated [G] and (2) designated [L], above, with just one class meeting per week. Each instructor involved taught all treatments, and all instructors had previous experience in both didactic and inquiry-based teaching.

Group work problems were created by a team of instructors and professors at UAB familiar with MA098. Problems were selected to be in line with upcoming, but not yet instructed on by the computer, material from the course. For instance, students working on familiarity with linear expressions might be asked to describe and come up with an algebraic rule for a growing pattern of toy blocks or garden stones that has been described for the first few stages of growth. Each problem was designed to include a challenge section at the end which allowed for further exploration by students who progressed more quickly through the problem, and some of the problems started with an also open-ended warm-up to stimulate group discussion. See Appendix below for examples of group work problems. Problems similar to the group work problems each week were presented by instructors in the lecture treatment classes.

Data gathered during the experiments in Fall 2009 and Fall 2010, and reported by Mayer (2010, 2011) on the two cohorts of elementary algebra students, included (1) course grades and test scores, (2) pre-test and post-test of content knowledge based upon a test which incorporated three open-ended problems, (3) for the 2010 cohort only, pre-test and post-test of content knowledge based upon a test consisting of 25 objective questions, (4) student course evaluations using the online IDEA system (IDEA 2010), and (5) RTOP observations of the instructors (RTOP 2010, Sawada 2002). For this study, (6) data on performance of students in the next mathematics course taken after the elementary algebra course and within the next three semesters was collected from the university data base.

Results of the Research

Analysis of student success in subsequent courses, as measured by students' final grade in the next course, was analyzed by using the comparisons of means independent t-test with an alpha of 0.05. Students' grades in subsequent courses were coded as follows: A-5, B-4, C-3, D-2, F-1. Figure 4 depicts statistics on students' grades for the Fall 2009 cohort in their subsequent math course making no distinction between subsequent courses. There was no significant difference between student grades in the next course based on the MA098 treatment (G or L) they received. Figure 5 breaks down the Fall 2009 cohort based on the specific subsequent course taken: MA110 is finite mathematics and is taught only in an inquiry-based/computer-assisted format and MA102 is Intermediate Algebra, taught only in a lecture/computer-assisted format. There was no significant difference between MA098 treatment groups for either MA110 or MA102 as the next course, though the MA098(L)→MA102 trajectory narrowly missed significance.

Fall 2009 Cohort				
Treatment	N	Mean	Standard Deviation	Significance (2-tailed)
Lecture (L)	132	3.6591	1.0101	0.244
Group (G)	129	3.5116	1.03166	

Figure 4. No distinction made between courses taken subsequently.

Fall 2009 Cohort					
Next Course	Treatment	N	Mean	Standard Deviation	Significance (2-tailed)
MA110	Lecture (L)	54	3.7407	0.91497	0.679
	Group (G)	56	3.8036	0.64441	
MA102	Lecture (L)	77	3.6234	1.06424	0.067
	Group (G)	72	3.2778	1.21287	

Figure 5. Distinguishing between MA102 (Intermediate Algebra) and MA110 (Finite Mathematics) taken subsequently.

There were three treatments in the Fall 2010 cohort: GG, LL, and GL. Figure 6 shows data on how these treatment groups compared pair-wise based on student success in subsequent courses, making no distinction between the next two possible courses. There were no significant differences between any of the three MA098 treatments as measured by final grades in subsequent courses.

Fall 2010 Cohort				
Treatment	N	Mean	Standard Deviation	Significance (2-tailed)
Lecture (LL)	79	3.5063	1.0484	0.913
Group/Lecture (GL)	76	3.5263	1.2052	
Lectures (LL)	79	3.5063	1.0484	0.323
Group (GG)	66	3.3333	1.0246	
Group/Lecture (GL)	76	3.5263	1.2052	0.308
Group (GG)	66	3.3333	1.0246	

Figure 6. No distinction made between courses taken subsequently.

Fall 2010 Cohort: Next Course MA110				
Treatment	N	Mean	Standard Deviation	Significance (2-tailed)
Lecture (LL)	29	3.9655	.9813	0.987
Group/Lecture (GL)	26	3.9615	.7736	
Lectures (LL)	29	3.9655	.9813	0.270
Group (GG)	24	3.7083	.6902	

Treatment	N	Mean	Standard Deviation	Significance (2-tailed)
Group/Lecture (GL)	26	3.9615	.7736	0.227
Group (GG)	24	3.7083	.6902	
Fall 2010 Cohort: Next Course MA102				
Treatment	N	Mean	Standard Deviation	Significance (2-tailed)
Lecture (LL)	50	3.2400	1.0012	0.757
Group/Lecture (GL)	48	3.3125	1.2908	
Lectures (LL)	50	3.2400	1.0012	0.758
Group (GG)	41	3.1707	1.1158	
Group/Lecture (GL)	48	3.3125	1.2908	0.580
Group (GG)	41	3.1707	1.1158	

Figure 7. Distinctions made between subsequent courses

Summary

We found no differences in success in subsequent courses ascribable to treatment in MA098. Two possible reasons for this are (1) the measure we are using for student success in their next course is student final grade, which is largely based on the computer instruction assessments, and (2) student pass rates for the next course are already so high (see Figure 8 below). We have compiled descriptive statistics for students who have passed through MA098 in the falls of 2007, 2008, 2009, and 2010 and how they score in the next course if it was taken within three semester of passing MA098. The results that follow are pass percentages (where a passing grade is an A, B, or C, as opposed to D, F, or W) of students' next course after MA098, first as a whole, and then split between the two courses they can choose from that follow MA098. MA102 is an intermediate algebra course for students who want to continue taking mathematics courses, and MA110 is a terminal finite math course for students who require only one university math credit.

Fall 098 cohort	**2007**	**2008**	**2009**	**2010**
pass% in next course	86.0%	86.4%	86.0%	84.6%
%passed MA102	78.7%	76.9%	73.8%	80.6%
%passed MA110	97.4%	91.7%	88.2%	96.6%

Figure 8. Pass rates for Fall MA098 cohorts in the next course.

As shown by the data, there has not been a significant change in students pass rates for their next course after MA098 over these four years. This is true looking at both the students and their next course as a whole and when splitting them up by the next course taken. This suggests that the major cause for pass rates going up was UAB's switch to computer assisted instruction in 2007. We also hypothesize, for students who pass MA098, that since student pass rates in their subsequent course are so high, between 84.6% and 86.4%, that it is difficult to detect differences from year to year. Figure 9 shows the overall student pass rates in Fall semesters over six years, pre-dating the switch to computer-assisted instruction. This includes students who placed directly into MA 102 or MA110. Students who have succeeded in MA098 (in 2007-2010) appear to have an advantage over students who place directly into the subsequent courses, but further study will be required to confirm this conjecture.

Fall Semester	2005	2006	2007	2008	2009	2010
%passed MA098	40%	54%	56%	61%	71%	66%
%passed MA102	34%	53%	62%	75%	81%	74%
%passed MA110	69%	63%	72%	89%	85%	89%

Figure 9. Students passing Fall semester courses.

Questions for Further Research/Analysis

Compare students pass rates in MA102 and MA110 based on how students got into the course. Did they: place into the course via math ACT or SAT scores, pass the UAB math entrance test, or did they take MA098 at UAB? Did treatments from the Fall2010 experiment have a statistically significant effect when compared to each of these subgroups?

Implications for Practice

We now teach all regular sections of elementary algebra following the blended treatment of the Fall 2010 experimental cohort: three class meetings weekly, one inquiry-based, one lecture, and one in the lab. We made our decision to change MA098 instruction prior to analyzing student success in subsequence courses based upon gains on open-ended problems and student satisfaction. Though students do not appear to have done better in terms of course grades nor success in subsequent courses, we see that they have made gains in communication and problem-solving. In the 2008 study of a finite mathematics course, we gave a longitudinal post-test (one year delayed) to a sample of students. The gains made by the group work treatment on the open-ended pre/post-test were maintained on the longitudinal post-test (see Figure 10).

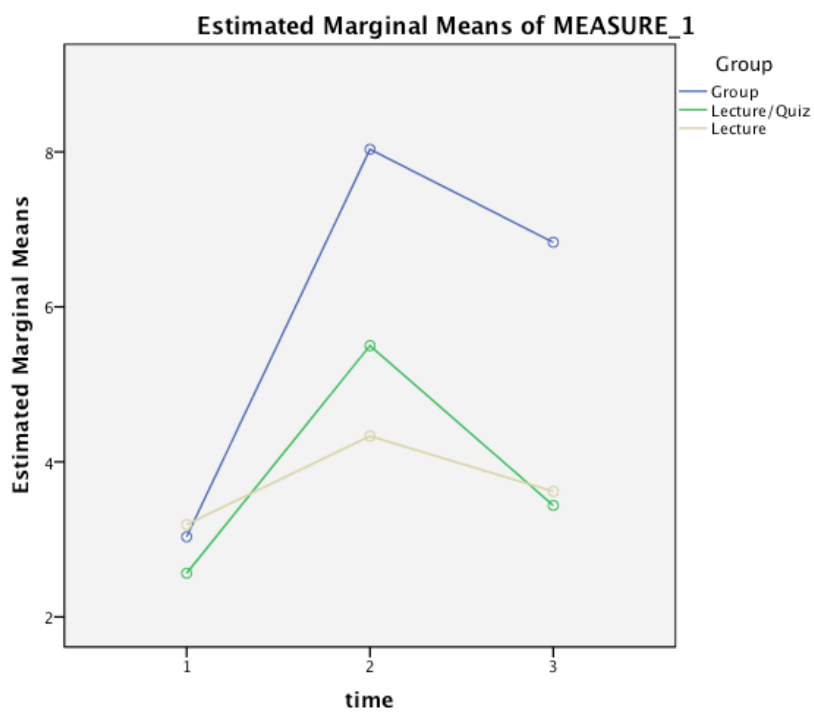


Figure 10. Open-Ended Pre/Post/Longitudinal Post-Test 2008 Cohort: N=67; Group=30; Lecture/Quiz=16; Lecture=21.

In view of the inherent coherence of algebra-related topics cutting across courses (Oehtrman, 2008), we expect to extend this study in subsequent years to credit courses such as intermediate algebra, pre-calculus algebra, and pre-calculus trigonometry, all of which presently incorporate computer-assisted instruction together with one weekly lecture meeting, and all in the course trajectory leading to calculus.

References

Alrø, H., & Skovsmose, O. (2002). *Dialogue and learning in mathematics education: Intention, reflection, critique*. Boston: Kluwer Academic Publishers.

Doorn, D. and O'Brien, M. (2007). Assessing the Gains from Concept Mapping in Introductory Statistics. *International Journal for the Scholarship of Teaching and Learning* 1(2); online at \url{http://academics.georgiasouthern.edu/ijsotl/v1n2/articles/doorn/index.htm}.

Duch, B. J., Groh, S. E., & Allen, D. E. (2001). *The power of problem-based learning.* Sterling, VA: Stylus.

Gautreau, R. and Novemsky, L. (1997). Concepts First – a Small Group Approach to Physics Learning. *Am. J. Phys.* 65(5): 418-429.

Gillies, R. M. (2007). *Cooperative learning: Integrating theory and practice*. Thousand Oaks, CA: Sage Publications.

Goodsell, A., Mahler, M., Tinto, V., Smith, B. L., & MacGregor, J. (Eds.) (1992). *Collaborative learning: A sourcebook for higher education.* University Park, PA: National Center on Postsecondary Teaching, Learning and Assessment.

Hoellwarth, C., Moelter, M. J., and Knight, R.D. (2005). A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms. *Am. J. Phys.* 73(5): 459-463.

Hough, S. (2010a). The effects of the use of inquiry-based learning in undergraduate mathematics on student outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Hough, S. (2010b). The effects of the use of inquiry-based learning in undergraduate mathematics on prospective teacher outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

IDEA (2010). The IDEA Center. \url{http://www.theideacenter.org/our-services/student-ratings-instruction/0030-idea-online}.

Johnson, D., Johnson, R., & Smith, K. (1998). *Active learning: Cooperation in the college classroom,* 2nd ed. Edina, MN: Interaction Book Co.

Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19-44). Palo Alto, CA: Greenwood.

Laursen, S., Hasai, M-L., Kogan, M., and Hunter, A-B., *Evaluation of the IBL Mathematics Project: Student and Instructor Outcomes of Inquiry-Based Learning in College Mathematics.* \url{http://www.colorado.edu/eer/research/steminquiry.html}.

Marrongelle, K. and Rasmussen, C. (2008). Meeting New Teaching Challenges: Teaching Strategies that Mediate Between All Lecture and All Student Discovery. In M. Carlson and C. Rasmussen (Eds.), *Making the Connection: Research to Practice in Undergraduate Mathematics*, pp. 167-178. Mathematical Association of America, Washington, DC.

Mayer, J., Cochran. R., Stansell, L.R., Land, H.A., Bond, W.O., Fulmore, J.S., & Argo, J.H. (2009). Incorporating Inquiry-Based Class Sessions with Computer Assisted Instruction. Electronic Proceedings of 2009 SIGMAA Conference on Research in Undergraduate Mathematics Education, Raleigh, NC, February, 2009.

Mayer, J., Cochran, R., Fulmore, J., Ingram, T., Stansell, L., Argo, J., and Bond, W. (2010). Blending Inquiry-Based and Computer Assisted Instruction in a Basic Algebra Course: a Quasi-Experimental Study, Electronic Proceedings of Conference on Research in Undergraduate Mathematics Education, Raleigh, NC, February, 2010.

Mayer, J.C., Cochran, R., Fulmore, J., Ingram, T., Stansell, L., and Bond, W. (2011), Inquiry-Based and Didactic Instruction in a Computer-Assisted Context. Electronic Proceedings of Conference on Research in Undergraduate Mathematics Education, Portland, OR. February, 2011.

Oerhtman, M., Carlson, M., and Thompson, P. (2008). Foundational Reasoning Abilities that Promote Coherence in Students' Function Understanding. In M. Carlson and C. Rasmussen (Eds.), *Making the Connection: Research to Practice in Undergraduate Mathematics,* pp. 65-80. Mathematical Association of America, Washington, DC.

Perkins, D. N., & Tishman, S. (2001). Dispositional aspects of intelligence. In Collis, J. M., & Messick, S. (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement* (pp. 233-257). Mahwah, NJ: Lawrence Erlbaum.

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.

RTOP (2010), Reformed Teaching Observation Protocol. Developed by ACEPT Project (Arizona State University). \url{http://physicsed.buffalostate.edu/AZTEC/RTOP/RTOP_full/}

Savin-Baden, M., & Major, C. H. (2004). Foundations of problem-based learning. Maidenhead, UK: Open University Press.

Sawada, D., Piburn, M., 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(6), 245-253.

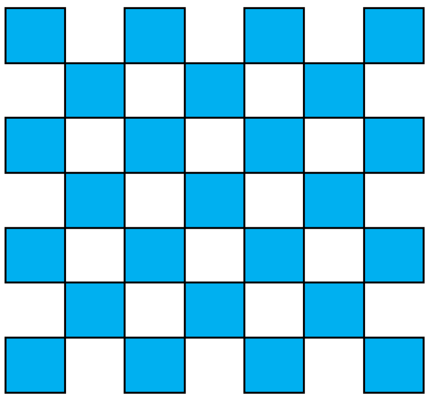
Appendix

Example Problem 1: A truck travels 180 miles on the highway in the same amount of time it travels 40 miles in the city. If the rate that the truck is traveling in the city is 30 mph slower than on the highway, find the rates at which the truck was driving both on the highway and in the city.

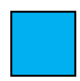
Some examples of student solution ideas:

1) After noticing that the time is the same, students set up an equation involving the ratios of the different distances and rates.
2) Students guess and check rates.
3) Students notice that going 30 miles an hour faster causes the highway trip to be able to cover an additional 80 miles.

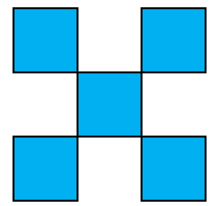
Example Problem 2:



Stage 4



Stage 1



Stage 2

Above are three stages in a growing pattern of square tiles. Build two more structures in the pattern. How many tiles will each take? How many tiles are needed for the 10th structure? Write an algebraic rule to find the number of tiles needed for any stage of growth. Define your variables. Show geometrically why your rule makes sense.

Some examples of student solution ideas:

1) Students make a table showing that the amount of tiles added from one stage to the next goes up by 4 each time. This usually leads to a recursive answer.
2) Students build the next stage from a preceding stage by putting tiles around the edge in four strips. The number of tiles added is then 4(stage number-1). From this, students can again come up with a recursive formula.
3) Students find a closed form solution based on the observation that in stage X there are X rows with X tiles and (X-1) rows with (X-1) tiles in them.
4) Students find a closed form solution based on the observation that in stage X there are X columns with X tiles and (X-1) columns with (X-1) tiles in them.
5) Students find a closed form solution by rearranging tiles to form two squares.
6) Students find the “area” of the whole figure, divide by 2, and round up.

DOES A STATEMENT OF WHETHER ORDER MATTERS IN COUNTING PROBLEMS AFFECT STUDENTS' STRATEGIES?

Todd CadwalladerOlsker, Nicole Engelke, Scott Annin, and Amanda Henning
California State University, Fullerton

Counting problems ask students to compute the number of ways a certain set of requirements can be satisfied, and they are important in such mathematical subjects as probability, combinatorics, and abstract algebra, among others. Students are often taught to solve counting problems by looking for specific clues to help categorize the problems and identify solution strategies. In this study, we investigate how the wording of certain counting problems, specifically whether or not "order matters," affects students' solution strategies. In particular, we gave students questions involving explicit statements as to whether or not order matters, some of which were intentionally misleading, and questions that do not contain such an explicit statement. Data was collected in the form of written responses and student interviews. The results show that many students do, in fact, rely heavily on such explicit statements about whether order matters, even when such statements are misleading.

Key words: combinatorics, permutations, combinations, problem solving

Introduction

Counting problems are a type of combinatorial problem which ask the solver to determine the number of ways a certain set of requirements can be satisfied in a given situation. For example, the problem might ask, "How many 5-card poker hands contain cards all the same suit?" Such questions arise in elementary probability questions in high school classes, in more advanced probability classes at the undergraduate level, as well as in abstract algebra, combinatorics, and other areas of the undergraduate curriculum.

Students are given several tools to solve counting problems. The two most basic tools are the multiplication principle (also known as the fundamental counting principle) and the addition principle. Students are also introduced to some useful formulas: the combination formula $C(n,k)$ counts the number of unordered subsets of size k that can be made from a set of size n; the permutation formula $P(n,k)$ counts the number of ordered subsets. Both of these formulas are derived from the multiplication principle, and can be viewed as "shortcuts" for specific applications of the multiplication principle. In almost every textbook used in the United States, these formulas are defined (as above) in terms of a selection model, in which a sample of elements is drawn from a set of objects. In some problems, repetition of selected objects is allowed. Therefore, four basic combinatorial operations can be defined as in Table 1 (Godino, Batanero, and Roa, 2005; Batanero, Navarro-Pelayo, and Godino, 1997; Rosen, 2011). We use the notation $P(n,k)$ for permutations without repetition, $C(n,k)$ for combinations without repetition, $PR(n,k)$ for permutations with repetition, and $CR(n,k)$ for combinations with repetition.

While other combinatorial models (distribution, partition) can appear in counting problems (Dubois, 1984, cited in Batanero, et. al. 1997), the selection model is the most familiar to most students, and solving problems using other models often involves translating the problem into a selection model (when possible) and applying one of the basic combinatorial operations (Godino et al., 2005). Several student difficulties with counting problems have been identified in the literature, and students may be more or less prone to make errors depending on several factors: the type of combinatorial operation (permutation or combination, with or without repetition); the nature of elements to be combined (letters, numbers, people, or objects); the implicit combinatorial model (selection, distribution, or partition); and the values given to n and k (Fischbein and Gazit, 1988; Batanero et al., 1997; Eizenberg and Zaslavsky, 2004).

Batanero et al. (1997) also catalogue several types of student error. In particular, one type of error is the "error of order," which Batanero et al. (1997) describe as, "confusing the criteria of combinations and arrangements, that is, distinguishing the order of the elements when it is irrelevant or, on the contrary, not considering the order when it is essential." This issue will be the focus of this study.

As noted earlier, most students are familiar with counting problems based on a selection model. Students are often taught to solve such problems by identifying the sampling conditions of the problem, recognizing the appropriate combinatorial operation (as in Table 1), and applying the required formula. While it is well-known that students often have difficulty recognizing the appropriate combinatorial operation (Batanero et al., 1997; Eizenberg and Zaslavsky, 2004; Godino et al., 2005), there have not been, to our knowledge, any studies examining the strategies students use to identify the combinatorial operation. Students are often taught to focus on whether or not order is allowed, and whether or not repetition is allowed. However, even in simple counting problems, these factors may not be obvious, and in fact, can be somewhat misleading. For example, consider the problem:

> A club has five members. In how many ways can a president, vice-president, and treasurer be elected?

The standard solution to this problem interprets this as "permutations without repetition," $P(5,3)$, assuming that a club member cannot simultaneously hold more than one office. However, it is not immediately clear to many students exactly how "order matters" in this problem. One explanation is that the selection of three officers can be mapped to an ordered subset of the club members by making the first selected member to be the president, the second to be the vice-president, and the third to be the treasurer. However, there are other ways in which the order does not matter: for example, the order in which the elections are held does not matter. Thus, in problems of this type, the question of whether or not "order matters" may not be the right question, and perhaps a different strategy might be more successful for students.

Theoretical Perspective

Batanero, Nevarro-Pelayo, and Godino (1997) draw on the work of Dubois (1984), classifying simple combinatorial ideas into three models. *Selections* are the most familiar to most students, and draw on the idea of selecting objects from a set. *Distributions* emphasize the concept of mapping a set to another set, and *partitions* divide a set into subsets. Because the selection model is the most familiar to students, this work will focus on that model. Within the selection model, Batanero et. al. (1997) identify four basic combinatorial operations: permutation[1], combination, permutation with repetition, and combination with repetition. These operations are defined in Table 1 below. These operations are often derived from the fundamental counting principle, and it is possible that students may employ the fundamental counting principle directly rather than applying one of the formulas in Table 1. By doing so, the students may be implicitly thinking of permutations, for example, as a process involving several operations rather than a single operation.

[1] Batanero et. al. (1997) and others, including Eizenberg and Zaslavski (2004) generally use the word *arrangement* rather than *permutation*, and reserve the term *permutation* to refer to arrangements of the entire set. We will use the term *permutation* to refer to both cases.

	Ordered Sample	Unordered sample
No repetition	□□,□ =□!□−□ !	□□,□ =□!□!□−□ !
Repetition	□□□,□ =□□	□□□,□ =(□+□−1)! □!□−1!

Table 1: Four basic combinatorial operations within the selection model.

Several student difficulties with counting problems have been identified in the literature, and students may be more or less prone to making errors depending on several factors: the type of combinatorial operation (permutation or combination, with or without repetition); the nature of elements to be combined (letters, numbers, people, or objects); the implicit combinatorial model (selection, distribution, or partition); and the values given to *n* and *k* (Batanero et. al. 1997, Eizenberg et. al. 2004).

It is well documented that students prefer to learn from examples (Bassok, Chase, & Martin, 1998; Bernardo, 2001; Campione, Brown, & Connell, 1989; Catrambone, 1994, 1995, 1996; Kulm & Days, 1979; Lithner, 2003; Robertson, 2000; Schoenfeld, 1989; Sierpinska, 1995; Silver & Marshall, 1989). While this may be preferred by students, evidence suggests that it is not as successful as one may hope (Bernardo, 2001; Kulm & Days, 1979; Robertson, 2000; Schoenfeld, 1989; Silver & Marshall, 1989). Silver and Marshall (1989) indicated that while experts focus on structure, novices tend to focus on superficial features when attempting to solve word problems resulting in poor transfer. Robertson (2000) indicated that mapping errors (not being able to properly choose and align similar problems), over-transfer/matching errors (not being able to properly assign values to variables in the problem) and frame errors (inability to generate and adapt the relevant equation) are extremely common. The use of examples to structure the problems solving process is currently being studied by some researchers such as Sinclair, Watson, Zazkis, and Mason (2011) as a personal example space.

As most counting problems involve context about what is to be counted, there is a natural connection between what strategies students choose to use and their personal example spaces (PES). Personal example spaces are defined to be the set of available examples and methods of example construction a learner has at their disposal for solving problems. The “order matters” heuristic approach to combination problems is an example of what we expect to be prevalent in students’ personal example spaces. Sinclair, et al (2011) addressed how personal example spaces are structured, paying particular attention to the varying degrees of “connectedness” such PESs may have. The more connected one’s example space, the greater the likelihood of having a stronger understanding of the concept. They also indicate that slightly different prompts may trigger the use of different examples.

We believe the PES strongly influences students’ strategies while solving counting problems in myriad ways. For example: 1) key phrases such as “order matters” triggers comparisons and strategies used in examples with similar phrasing, 2) similar context triggers the use of the exact same strategy (see the stuffed animal example in the results), and 3) more examples complicate or confuse what is known.

Methods

We claim that the burden for successfully answering questions about combinations and permutations often falls upon the solver’s careful reading and interpretation of the problem. In particular, we believe that a student’s interpretation of whether order matters, and what it means for order to matter, greatly impacts that student’s thinking. We believe that the wording of questions in this area has a crucial impact on how students view them.

To investigate these claims, we prepared a written quiz of six combinatorics problems; see Appendix. Problems 1, 3, 4 and 5 make statements concerning whether or not “order matters,” with Problems 3 and 5 written intentionally to present the question of “order” in a non-standard way. As our results will show, these statements may have influenced students to solve those problems incorrectly.

This quiz was given to students enrolled in a combinatorics course and discrete math course at a large state university during the Fall 2011 semester. There were thirteen graduate students and twenty-one undergraduate students who participated in this study. This quiz was given twice, once before the students received direct instruction about combinations and permutations and again a few weeks after. We will often refer to these two rounds of the quiz as the pre-quiz and the post-quiz.

A group of ten graduate students were also interviewed following both rounds of the quiz regarding their thinking process on the quiz. The interviews were video recorded and analyzed, and pseudonyms were assigned to each student.

Direct instruction in the two classes differed slightly. In the combinatorics class, strategies for solving combinatorics problems was the second topic covered, following the Pigeonhole Principle. The topic began by covering the sum and multiplication principles for counting, describing these rules in terms of tasks to be completed, along with a basic example of each rule.

Next, the instructor covered permutations of sets and gave several basic examples. While he did introduce the notation *P(n,k)*, he rarely used it in practice, preferring instead to use factorial notation. In solving problems such as this where order matters, he drew a visual consisting of a sequence of “slots,” and he placed the number of choices for the task associated with each slot in that slot’s position. Next, he covered combinations of sets. He used the notation *C(n,k)*, and in this case, he tended to use this notation more consistently (rather than converting to factorials again). He showed the mathematical relationship between *C(n,k)* and *P(n,k)* and then gave a few other basic examples involving *C(n,k)* dealing with arrangements of men/women, cards from a deck, committee formations, license plates, etc.

The instructor concluded by pointing out that most counting problems are not so clear-cut. It is possible for the “order of objects” to somewhat matter and somewhat not matter. The typical example of this is the number of arrangements of the letters of MISSISSIPPI. Finally, the instructor provided a group work activity involving several variations on a counting problem, in order to impress upon students how slight variations in counting problems can lead to completely different analysis, tools, and answers.

In the discrete mathematics class, counting problems were not covered until the fifth week of class. Discussion of counting problems also began with the sum and multiplication principles. Examples and explanations based on tree diagrams were provided. Combinations and permutations were introduced as “shortcuts” for the multiplication principle, based both on a determination of whether or not order is counted, as well as an alternative way of thinking based on the idea of labeling. This class used the binomial coefficient notation for combinations, and used no special notation for permutations in class. However, the textbook used the notation *P(n,k)* for permutations.

The instructor discussed some of the misconceptions that can arise from focusing on whether or not order matters, and encouraged students not to focus on order. Rather, the instructor asked students to think of choosing a subset from a set, and to determine whether the chosen elements needed to be “labeled” or not. The subset can be chosen using combinations, as usual. If the chosen elements need to be labeled, then there are □! ways to apply those labels. This corresponds to permutations, and □□ · □! reduces to the standard permutation formula.

After discussing standard combinations and permutations, variations involving repetition or indistinguishable objects were introduced. The usual formulas for combinations with repetition allowed and permutations with repetition allowed were discussed in the usual way, as well as using the idea of labeling.

Results

Data was collected both before direct instruction and afterwards, in both classes. We tallied the number of correct and incorrect responses to each problem on the quiz and conducted interviews with the group of graduate students from the discrete mathematics class. Data analysis was conducted with attention to how students interpreted the questions, particularly in regard to phrasing about whether or not order matters. Problems 3 and 5 were written specifically to "mislead" students by including a statement about order that does not conform to the usual meaning of whether or not order matters.

Table 2 shows the number of correct and incorrect responses to the quiz, with the questions separated into three categories. Problems 1 and 4 contained a straightforward statement of "order matters" or "order does not matter," without trying to mislead the students. These questions are called positive in Table 2. Problems 3 and 5 contained a misleading statement about order, and are called negative in Table 2. Problems 2 and 6 did not contain a statement about order, and are called neutral in Table 2.

Pre-Instruction	Correct	Incorrect	% Correct
Positive	40	22	64.5%
Negative	14	48	22.6%
Neutral	41	21	66.1%

Post-Instruction	Correct	Incorrect	% Correct
Positive	52	10	83.9%
Negative	30	32	48.4%
Neutral	41	21	66.1%

Table 2: Counting problems presented to student participants

Tables 3 and 4 in the appendix show a summary for pre- and post- instruction interviews, making note of any instance in which the interviewee used "labels" (or equivalently, "slots"), used one operation or several, or made mention of "order." The charts are labeled green if the solution was ultimately correct, red if incorrect, and yellow if only a minor error was made (such as miswriting a number). We discuss the results of the interviews, both pre- and post-instruction, below.

Pre-Instruction Interviews

We hypothesized that the phrasing "order does not matter" and "order matters," particularly in Problems 3 and 5, would result in students identifying those key phrases and using a combination or permutation formula accordingly. Several students did just that. In her interview, Jane summarized her strategies with, "I remember in high school learning about if order matters, it is a permutation, and if order does not matter, it is a combination."

Problem 3 stated: A youth hockey team has twelve members. How many ways are there to choose a starting lineup of center, left wing, right wing, left defense, right defense, and goalie, if the order in which these positions are filled does not matter? This problem included the statement that "the order in which these positions are filled does not matter," which is true: the starting lineup is not changed if the position of goalie is filled before that of center, or vice-versa. However, this is not the standard meaning of “order does not matter.” In fact, since the order in which the positions are filled does not matter, the usual approach to this problem is to define an arbitrary order of positions (center first, left wing second, etc.) and map ordered subsets of players to starting lineups using this order of positions. By focusing the attention on the order of the positions, the problem misled students who relied on the “order matters” principle.

For some students, the phrase “the order in which these positions are filled does not matter” was an indication to use the choose function, □□. In the written results, eighteen of the participants used some form of a combination formula in their response, while only seven participants correctly solved the problem.

In her interview, Fiona stated, “When I saw the words ‘does not matter,’ I said OK that’s a choosing problem, I remember 12 choose 6 is the formula where order wouldn’t matter.” Similarly, Ruth’s response was, “This problem is almost the same as the first problem, except order does not matter. Order does not matter is combinations and order matters is permutations.” Another student, Emily stated, “It said the order does not matter, and so I know that we are choosing positions and we are going to divide by the number of positions factorial because order does not matter, so we are going to take away the redundant orders.” When prompted by the interviewer about the phrasing of the problem, she elaborated that without the words “order does not matter” she would have used a different strategy.

Other students were not swayed by the phrasing of the problem. Lowell stated, “It says that the order that the positions are chosen does not matter, which made it sound like it was going to be a combination problem instead of permutations. But each of the positions is different so it matters which person gets chosen in which position, so the order that you pick them does still matter, which put it back in a permutations question.”

This problem was functionally identical to Problem 6, which states that all colors on the main, trim, accent and siding must be different, but makes no explicit statement of whether or not order matters. This change in wording made it much clearer to the students that a combination formula was unnecessary here, and more students were successful in solving Problem 6 than Problem 3.

Problem 5 stated: A toddler has an essentially unlimited supply of red and blue blocks, and is building stacks of these blocks. If the toddler makes a stack of eight blocks, how many ways are there to stack the blocks so that exactly three blocks are red? (The order in which the blocks are stacked matters.)

This problem included the statement “the order in which the blocks are stacked matters.” This statement is not misleading on the surface: outcomes of stacked blocks are different if the same blocks are rearranged, and this problem can be solved as permutations with repetition allowed. However, a simple permutation formula (without repetition) cannot be applied. In fact, one way to solve the problem is to choose an (unordered) subset of the eight positions to be filled with red blocks, leaving the remaining five positions to be filled with blue blocks. That is, even though “order matters,” a combination formula can be appropriately used to solve this problem.

Fewer students behaved according to our hypothesis on this problem: Five students clearly indicated the use of a permutation formula in their written work. However, this was also the most difficult problem: only six participants gave a completely correct answer. None of the interviewed students approached the problem as a permutations with repetitions allowed problem, and if they did attempt to use a permutation formula, it was for permutations without repetition. Jane, using her primary strategy (stated above) said, “Now since the order does matter, because we are thinking about lining things up, that’s why I use a permutation.”

Emily, on the other hand, wrote $8!3!5!+8!5!3!$, and described her solution by "choosing" the three blocks to be red and "choosing" the five blocks to be blue. Here, she discarded her strategy of using the key phrase "order matters" and used combination formulas (in conjunction with the addition principle, in the mistaken belief that both terms were needed to account for the two colors). Emily indicated that the fact that the blocks were indistinguishable prompted her to modify her strategy of focusing on whether or not order mattered. During her interview, she used the term "choose" to refer to any kind of selection process, with or without order (or in terms of the multiplication principle), but also distinguished (with less than total confidence) between "*n* choose *k*" and "*n* factorial over *k* factorial" (a misremembered version of □!□−□ !).

Lowell, again undisturbed by the statement about order, described his strategy as looking at the positions of the blocks and "choosing" three of them to be red. While Lowell was very successful in avoiding this pitfall, it should be noted that Lowell was simultaneously enrolled in both classes involved in this study, and therefore had much more recent experience with counting problems.

This problem also produced an unanticipated phrasing difficulty for students: at least three students interpreted the requirement that "exactly three blocks are red" to mean that the three red blocks were to be stacked adjacently.

Post-Instruction Interviews

In the interviews following instruction, many of the themes that emerged during the interviews following the pre-quiz were present again. In general, the students felt more confident about their answers than they did after the pre-quiz. Most students' impressions that "order matters" necessitate a permutation approach while "order does not matter" necessitates a combination approach continued on the post-quiz. In response to Problem 1, for example, Emily stated "the order mattered so then I figured this is a permutation" and wrote the word "permutation" on her solution. By contrast, on Problem 3, she was misled by the words "order does not matter" when she said "The order doesn't matter, so picking each person is not going to affect which position…..so that's how I got the combination instead of a permutation," and wrote the word "combination" on her solution. (See Appendix for Emily's written work.) When asked if she would have answered the problem differently if the phrase "order does not matter" had not been there, she said "Yes, I would have assumed that each choice was made for certain positions." Nonetheless, Emily stated she was "very confident" in her answer to this problem. Finally, on Problem 5, Emily used a combination to solve the problem and the interviewer questioned her about his by asking: "There's a statement that says the order matters….normally that means a permutation, like earlier, but here you used a combination….," to which she replied "It's because the blocks are indistinguishable…, like picking the first, second, and fourth to be red is the same as picking the fourth, second, and first…"

On Problem 5, Jane (whose answer given in the form of a combination was correct) explained in her post-instruction interview that she was not confident in her answer because "Now I'm thinking of it, the order does matter so it should be a permutation instead of a combination." As a result, Jane changed her answer.

Lowell is the only student interviewed who got all six questions correct on the post-quiz. He talked in his interview about comparing the quiz problems with techniques learned in class, and seeing and remembering tricks. While he found the wording on some questions misleading, he ultimately managed to arrive at the correct answers. On Problem 3, Lowell indicated that the problem "gave me pause for awhile the first time around because of the very last sentence where it says the order that the positions are filled does not matter, because normally that's the keyword that tells you that you're doing a combination." He continued to describe a labeling strategy described in his Discrete Math class to facilitate his understanding of why the order does matter in this problem: "When we talked about order mattering or not mattering, one of the ways was… applying labels to them. So instead of the order being how you're placing them, you're applying a label and that's order mattering or not. So, in this one there were already labels for the positions,… so that was counting as an order. So picking them in a specific order, that was just kind of an extra piece that isn't normally what gets counted."

Likewise, on Problem 5, Lowell recognized that the phrase "order matters" is misleading and gave a combination form for his answer, although he felt it was "trickier" than Problem 4 because of the words "order matters."

While students felt that the quiz was easier following instruction on the material, there is evidence that some students may have gotten confused in trying to make comparisons between examples shown in class and problems on the quiz. At the outset of the post-instruction interview, Allison stated that she "felt less confident on the post-quiz than the pre-quiz. This was at odds with most students' opinions. She said "I was trying to compare them to examples we had seen in class and trying to think if it was like this problem or that problem." Later in the interview, she said "I think doing a lot of examples kind of can be a good thing and kind of can be a bad thing because….you can get confused by a lot of things. Another thing that happened in class is that every time we got an example, people would be like `What if we changed it slightly this way' and then it changes the whole problem. It's really hard to follow a formula of some sort." In many early math classes, having numerous examples helps students internalize computational procedures. To emphasize Allison's point, slight changes in the wording of combinatorics problems frequently lead to a different strategy requiring the application of different concepts. Hence, echoing what was seen by Sinclair et al. (2011), having numerous loosely connected examples from which to base a solution strategy may lead to less success in the problem solving process.

An excellent example of the challenge Allison described came out during Ruth's interview where she tried to apply a formula for a technique described in class as stars-and-bars to solve Problems 3 and 6 on the post-quiz. In explaining Problem 3, Ruth said "I thought this might be like what he taught with the stars and bars [combination with repetition]. The question we did in class had to do with different types of doughnuts….it's up to you how many you pick from each category." She went on to say "I feel a little better about that one." We hypothesize that Ruth's choice of example on which she based her solution strategy came from the manner in which she chose to identify relevant aspects of the problem statement. For Problem 3, she wrote "C | LW | RW | LD | RD | G" and on Problem 6 she wrote "MC | TC | AC | SC." This notation is commonly used to set up the prototypical example about choosing a dozen doughnuts from a specified number of types of doughnuts and then generalized using stars and bars. Here, Ruth is conflating the ideas of choosing doughnuts from a given number of types with the idea of choosing a particular player to fill a position. In the doughnut problem, all doughnuts of a particular type are isomorphic; in contrast, choosing a particular player for a position makes a big difference in the lineup. Hence, Ruth's choice of notation causes her to choose an example from her PES that does not accurately capture the salient aspects of the problem.

In Madeline's interview, she indicated that she second-guesses herself a lot while trying to draw comparisons between problems. For instance, she was not confident in her answer to Problem 4; she had seen a similar problem on a test in her class: "There was a stuffed animal problem on the test, and I bombed it, so I wasn't so sure about this one….I'm just trying to remember how the test was. This one looks more difficult, in a way." She went on to say "I know it's just a simple problem, but maybe I was feeling like `Am I missing something here?' or 'Could it be different from #3?'" In this case, we believe that Madeline is focused on the context of choosing stuffed animals as an indicator to which strategy (one that she does not remember) should be chosen. However, these doubts do not dissuade her from using this strategy and notes that it makes certain problems easier.

In Problem 5, Madeline indicated that she was using the bit string example from class to solve the problem. She noted, "I feel like this one is not as difficult. I feel like I have a problem that I can base it on. I always try to refer what I learn from the class and it looks similar to the bit string problem so I approach it that way." The particular "bit string problem" Madeline referred to is not quite clear, but the class homework included several problems that asked for the number of possible bits strings (consisting of 0's and 1's) of a certain length, containing certain patterns. She draws the parallel that in the bit string problem the choice was either 0 or 1, but here it is red or blue. On her paper, she has a few examples of possible towers in which all the reds are grouped together and shifting by one position in each exemplar. In one bit string problem done in class, the bit strings required the pattern "000;" Madeline may have been thinking of this pattern when creating her towers. As a result, Madeline makes a partially correct connection between the current task and a known example.

During Fiona's post-instruction interview, she made explicit reference to similar problems shown in class on four of the six questions on the quiz (Problems 1, 3, 4, and 5). Further, she also indicated that no problem similar to Problem 2 had been covered in class. In contrast to Ruth and Madeline, Fiona was very successful at using her PES to solve problems. Our evidence suggests that students' use of PESs to choose solution strategies is common and can have mixed consequences on the students' success in solving problems.

Conclusion

Counting problems can be quite difficult, and many different types of error are possible. Our study takes a closer look at one dimension of the error types identified by Batanero et al. (1997); namely, that of the "error of order." To help students avoid this error, instructors and textbooks have adopted a single organizing principle for dealing with combinations and permutations: "If order matters, use permutations; if not, use combinations." However, such a principle belies the difficulty of such problems, and in fact, can be misleading.

In our two misleading problems, we gave statements about order that do not conform to the usual meaning of "order matters." The usual interpretation of "order matters" is that, when a subset is selected from a set, a difference in the order in which the elements of the subsets are selected constitutes a different outcome. In other words, ordered subsets are counted.

In answer to the title question of this paper, "Does a statement of whether order matters in counting problems affect students' strategies?," the participants in our study do, in fact, seem to have been influenced by statements of order. Jane's post-instruction interview response to Problem 5 is particularly telling: she changed her correct response to an incorrect response based on the statement that "the order in which the blocks are stacked matters." While some students (particularly Lowell) were able to devise strategies to interpret such statements, many students were unsuccessful in doing so.

We also note that when examining students' responses to problems that can be solved using the permutation formula (Problems 1, 3, and 6), students tended to use several tasks, linked by the multiplication principle, rather than the single task of applying the permutation formula itself. When responding to Problem 6, for example, most of the students who answered correctly gave an answer of $14 \cdot 13 \cdot 12 \cdot 11$, rather than $P(14,4)$. In their interviews, participants explained their answer by noting the number of choices available for each part of the house. Some participants then rewrote their solution as *14!10!*, but the majority did not do so. Allison made use of the phrase "choosing one at a time" to explain her use of several tasks in this way, and in contrast, used the phrase "choosing all at once" to denote the single task of using the choose function.

This research study has drawn attention to a number of issues relevant for instructors of combinatorics. Perhaps more than in other mathematical subjects students are exposed to, the wording of combinatorics problems is extremely delicate. On the two problems from the quiz in which the wording was intentionally misleading (Problems 3 and 5), our quantitative data shows that student performance was poorest. The source of the students' difficulty with these problems, the misleading wording, was explicitly identified in the qualitative data (both pre-quiz and post-quiz) by student remarks about how the wording impacted their strategies. The trigger that the phrase "order matters" requires use of a permutation and the phrase "order doesn't matter" requires the use of a combination is firmly engrained in many combinatorics students' minds. Even after direct instruction was provided in which students were cautioned to avoid relying too heavily on this heuristic, many students still faltered on the post-quiz.

While the students generally did better on the post-quiz than the pre-quiz on "positive" and "negative" problems, and just as well on "neutral" problems, direct instruction does leave some negative consequences in its wake. During the post-instruction interviews, some students indicated a tendency to try building associations between quiz problems and examples seen in class, even when such associations were either non-existent or superficial at best. This led some students to second guess the strategies they were using. Some students tried to apply a strategy learned in class incorrectly, simply because they believed a particular quiz problem to be similar to one they had seen previously in which the strategy was applied.

It is somewhat surprising and disappointing that participants did no better on the "neutral" questions post-instruction. These questions, in which no statement about order is given, are precisely the kind of problems we want our students to be able to solve.

Avenues for Further Research

A number of further questions have been fostered through this research study. As our data suggests, some students try to view the counting problems as a series of tasks that are to be enumerated separately and then combined with the multiplication principle. This is especially common when a permutation is used, a situation where students in the study rarely used the $P(n,k)$ notation. Problems 1, 3, and 6, which could all be successfully solved using the permutation formula, were most often successfully solved using the viewpoint of several tasks. A much smaller group of students successfully solved the same problems using the viewpoint of only one or two tasks.

As noted above, some students developed a strategy of trying to associate quiz problems with previously seen examples. As a result, we would be interested to know which examples are appropriated by students in to their PESs and which of these examples provide students the most leverage in terms of problem solving strategies.

Some students felt intimidated by the nature of the counting problems in that very slight changes to the wording or the problem itself lead them to an entirely different strategy. It would be interesting to investigate this phenomenon and, if proven to be an impediment to student learning, supply teachers with tools to mitigate this threat to students' confidence.

Finally, it would be of interest to look for a correlation between student success on counting problems and their confidence level in solving counting problems. Interestingly, while the majority of students felt more confident on the post-quiz than the pre-quiz, the many errors committed on the post-quiz reveal in some cases that this confidence was unfounded. Indeed, while students may have tended to be more confident about the problems that they had in fact solved correctly, many of them also felt the same way about problems they had not solved correctly! An understanding of any correlation that exists here would provide researchers with motivation to look for cues to provide students to help them critique and evaluate their own answers, thereby enhancing student intuition about whether they are or are not solving counting problems correctly.

References

Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology, 35*, 99-134.

Batanero, C., Navarro-Pelayo, V., and Godino, J. D. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics*, 32(2):181–199.

Bernardo, A. B. I. (2001). Analogical problem construction and transfer in mathematical problem solving. *Educational Psychology*, 21(2), 137-150.

Campione, J. C., Brown, A. L., & Connell, M. L. (1989). Metacognition: On the importance of understanding what you are doing. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of problem solving* (Vol. 3, pp. 93-114): Lawrence Erlbaum Associates.

Catrambone, R. (1994). Improving examples to improve transfer to novel problems. *Memory & Cognition, 22*(5), 606-615.

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87*(1), 5-17.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(4), 1020-1031.

Dubois, J. G. (1984). Une systematique des configurations combinatoires simples. *Educational Studies in Mathematics*, 15(1):37–57.

Eizenberg, M. M. and Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning*, 6(1):15–36.

Fischbein, E. and Gazit, A. (1988). The combinatorial solving capacity in children and adolescents. Zentralblatt fuer Didaktitk der Mathematik, 5:193–198.

Godino, J. D., Batanero, C., and Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics*, 60:3–36.

Kulm, G., & Days, H. (1979). Information transfer in solving problems. *Journal for Research in Mathematics Education*, 94-102.

Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational Studies in Mathematics, 52*, 29-55.

Robertson, I. (2000). Imitative problem solving: Why transfer of learning often fails to occur. *Instructional Science, 28*, 263-289.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Rosen, K. H. (2011). Discrete Mathematics and its Applications (7th Edition). McGraw-Hill, New York.

Schoenfeld, A. H. (1989). Problem solving in context(s). In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of problem solving* (Vol. 3, pp. 82-92): Lawrence Erlbaum Associates.

Sierpinska, A. (1995). Mathematics: "in context", "pure", or "with applications"? *For the Learning of Mathematics*, 15(1), 2-15.

Silver, E. A., & Marshall, S. P. (1989). Mathematical and scientific problem solving: Findings, issues and instructional implications. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 265-290). Hillsdale, N.J.: Erlbaum.

Sinclair, N., Watson, A., Zazkis, R., and Mason, J. (2011). The structuring of personal example spaces. *Journal of Mathematical Behavior*, 30, 291-303.

Vinner, S., & Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for Research in Mathematics Education*, 20, 356-366.

Appendix

Quiz Problems

(1) A bag contains 26 marbles, labeled A through Z. In how many ways can six marbles be chosen, where each of the six chosen marbles is different and the order in which they are chosen matters?

(2) A chess club has 9 members. If the club puts on a friendly tournament in which each member plays every other member exactly once, how many games will be played?

(3) A youth hockey team has twelve members. How many ways are there to choose a starting lineup of center, left wing, right wing, left defense, right defense, and goalie, if the order in which these positions are filled does not matter?

(4) A child has 8 different stuffed animals. When leaving to visit her grandmother, the child is allowed to select three animals to take along. How many ways are there for the child to select the three animals? The order in which the animals are selected does not matter.

(5) A toddler has an essentially unlimited supply of red and blue blocks, and is building stacks of these blocks. If the toddler makes a stack of eight blocks, how many ways are there to stack the blocks so that exactly three blocks are red? (The order in which the blocks are stacked matters.)

(6) A painter has fourteen colors of paint available. When painting a house, she needs to choose a main color, trim color, accent color, and siding color, and all of these colors must be different from one another. How many ways are there for the painter to pick colors for the house?

Quiz results pre- and post-instruction

NAME	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6
ALLISON	SLOTS SEVERAL TASKS	SEVERAL TASKS	SLOTS SEVERAL TASKS	ONE TASK	SEVERAL TASKS	SLOTS SEVERAL TASKS
DAVE	SEVERAL TASKS ORDER MATTERS	ONE TASK	SEVERAL TASKS	ONE TASK	?	SEVERAL TASKS
EMILY	ONE TASK ORDER MATTERS	SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	ONE TASK ORDER DOESN'T MATTER	ONE TASK	SEVERAL TASKS ORDER MATTERS
FIONA	SLOTS SEVERAL OPERATIONS ORDER MATTERS	SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	ORDER DOESN'T MATTER	ONE TASK	SEVERAL OPERATIONS ORDER MATTERS
HELEN	ONE TASK	ONE TASK	ONE TASK	ONE TASK	N/A	N/A
JANE	SLOTS ORDER MATTERS	SEVERAL TASKS	SLOTS SEVERAL TASKS ORDER-"I DON'T SEE A DIFF	ONE TASK ORDER DOES NOT MATTER	ONE TASK ORDER MATTERS	SLOTS SEVERAL OPERATIONS
JACK	ONE TASK	SEVERAL TASKS	TWO TASKS ORDER OF WAYS TO ARRANGE	BRUTE FORCE	BRUTE FORCE	TWO TASKS ORDER OF WAYS TO ARRANGE
LOWELL	ONE TASK ORDER MATTERS	SEVERAL TASKS	SEVERAL TASKS ORDER DOES STILL MATTER	ONE TASK ORDER DOESN'T MATTER	ONE TASK ORDER ENDS UP BEING POSITIONS	SEVERAL TASKS
MADELINE	SLOTS SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS	SLOTS SEVERAL TASKS ORDER?	SLOTS SEVERAL TASKS	SLOTS SEVERAL TASKS	SLOTS SEVERAL TASKS ORDER?
RUTH	SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS	SEVERAL TASKS ORDER DOES NOT MATTER	SEVERAL TASKS ORDER DOES NOT MATTER	SEVERAL TASKS	SEVERAL TASKS

Table 3: Pre-instruction interview results.

NAME	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6
ALLISON	SLOTS SEVERAL TASKS	SEVERAL TASKS	SLOTS SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	SEVERAL TASKS	SLOTS ONE TASK
DAVE	SEVERAL TASKS ORDER MATTERS	ONE TASK	ONE TASK ORDER DOESN'T MATTER	ONE TASK	SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS ORDER MATTERS
EMILY	SLOTS SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	ONE TASK ORDER DOESN'T MATTER	ONE TASK	SEVERAL TASKS
FIONA	SLOTS SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS	LABEL / SLOTS SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	SLOTS ONE TASK	ONE TASK ORDER DOESN'T MATTER
HELEN	ONE TASK	N/A	SEVERAL TASKS	ONE TASK	N/A	ONE TASK
JANE	SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	ONE TASK ORDER DOESN'T MATTER	ONE TASK	SEVERAL TASKS
JACK	ONE AND SEVERAL TASKS ORDER MATTERS	SEVERAL TASKS	SLOTS ONE AND SEVERAL TASKS ORDER MATTERS	ONE AND SEVERAL TASKS ORDER DOESN'T MATTER	SLOTS ONE TASK	ONE AND SEVERAL TASKS ORDER MATTERS
LOWELL	ONE TASK ORDER MATTERS	ONE TASK	LABELS ONE TASK ORDER MATTERS	ONE TASK ORDER DOESN'T MATTER	ONE TASK ORDER DOESN'T MATTER	LABELS ONE TASK ORDER MATTERS
MADELINE	SLOTS SEVERAL TASKS	SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER	ONE TASK ORDER DOESN'T MATTER	SEVERAL TASKS	ONE TASK ORDER DOESN'T MATTER
RUTH	SLOTS SEVERAL TASKS	SEVERAL TASKS	ONE TASK	SEVERAL TASKS	SLOTS ONE TASK	ONE TASK

Table 4: Post-instruction interview results.

Emily's post-instruction responses to Problems 1 and 3

1. A bag contains 26 marbles, labeled A through Z. In how many ways can six marbles be chosen, where each of the six chosen marbles is different and the order in which they are chosen matters?

permutation

$\underline{26} \cdot \underline{25} \cdot \underline{24} \cdot \underline{23} \cdot \underline{22} \cdot \underline{21}$

$= \frac{26!}{20!}$

1. $\frac{26!}{20!}$ ways

3. A youth hockey team has twelve members. How many ways are there to choose a starting lineup of center, left wing, right wing, left defense, right defense, and goalie, if the order in which these positions are filled does not matter?

6

combination

$\binom{12}{6} = \frac{12!}{6!6!}$

distinct

3. $\frac{12!}{6!6!}$ ways

A STUDY OF ABSTRACT ALGEBRA TEXTBOOKS

Mindy Capaldi
Valparaiso University

This study uses reader-oriented theory and the analysis of example spaces to understand abstract algebra textbooks. Textbooks can lay the foundation for a course and greatly influence student understanding of the material. Multiple undergraduate abstract algebra texts were studied to investigate potential audiences of the books, the level of detail in explanations, examples, and proofs, and the overall material included in the book. Conclusions were drawn regarding some discrepancies between the intended reader and the implied reader and the appropriateness and differences among example spaces.

Key words: [Textbooks, Abstract Algebra, Reader-Oriented Theory, Example Spaces]

Introduction

Although there has been some significant research on mathematics textbooks, much of it has focused on the K-12 level (K-12 Mathematics Curriculum Center, 2005). The calculus reform movement motivated an extension of the study of textbooks into the collegiate level, but still the focus remained on the design and role of lower-level mathematics or calculus books. Little work has been done to investigate the use, purpose, strengths, and disadvantages of upper-level mathematics textbooks, especially for an abstract algebra course. Often these higher-level courses are where even mathematically gifted students start to struggle, leading to the need for research on how their learning experience can be improved. Many teachers, even in abstract algebra, use the textbook as a foundation, if not an outline, of the course material. As Robitaille and Travers (1992, p. 706) stated, “Teachers of mathematics in all countries rely heavily on textbooks in their day-to-day teaching, and this is perhaps more characteristic of the teaching of mathematics than of any other subject in the curriculum. Teachers decide what to teach, how to teach it, and what sorts of exercises to assign to their students largely on the basis of what is contained in the textbook authorized for their course.” It is important to understand what teachers are doing in the classroom, but because the source of methods and topics for many teachers stems from the course textbook, it is equally important to study the text.

Authors, even within the field of undergraduate abstract algebra textbooks, have different intentions for the content and use of their texts. Also, generational differences on how mathematics should be presented and learned can affect the language and style of the text. The range of popular abstract algebra textbooks today still includes many that were first written decades ago. Of course, these have usually been updated with new editions, but in general the changes from one edition to the next are not substantial. Modern theories of learning indicate the need for student-oriented teaching methods and reader-oriented textbook methods (Weinberg & Wiesner, 2011). Teachers, and textbooks, are no longer meant to simply “cover” material, but should facilitate a learning environment that inspires curiosity, speculation, inference, and quantitative literacy. Student thinking, and the multiple strategies that it may involve, should be valued (Reys, Reys, & Chaves, 2004).

The following article evaluates several popular and respected abstract algebra textbooks through the lens of reader-oriented theory and an analysis of example spaces. These two aspects of the study merge to create a new framework that allows one to make conclusions about whether the texts are successful and fit into modern ideas of learning. The methods employed in this study can be extended to other abstract texts or upper-level mathematics textbooks, and can aid teachers who are choosing a text for their class, or authors who are writing a new textbook.

Literature Review

1.1 What is reader-oriented theory?

Rosenblatt first proposed "reader-response theory" as an investigation of the reader's relationship with books, geared toward the teaching of English literature. She distinguished between two ways of experiencing a text, the *efferent* and *aesthetic*. Efferent readers look to acquire information and simply understand what the text is saying. Aesthetic readers' primary goal is to impose their own unique, personal engagement with a text through remembrances or opinions (Rosenblatt, 1938). Even after her work, research on pre-collegiate textbooks focused on reading as a relatively non-interactive experience, wherein the reader simply extracts information and duplicates it, departing from this idea of reader-response. Some science educators have used reader-oriented theory in evaluating student use of science textbooks, but it was not applied by mathematics educators in particular until recently (Weinberg & Wiesner, 2011).

Weinberg and Wiesner took reader-oriented theory into the realm of mathematics in an effort to "describe the characteristics of textbooks and readers that influence the ways students use textbooks to learn mathematics" (2011, p. 50). Within this theoretical framework, the use of textbooks moves beyond considering them as a static collection of ideas from which meaning is removed, and instead considers a student's active engagement with the material and the processes of reading and understanding. In other words, "the meaning of a text does not reside in the text itself, but rather is generated through a transaction between the text and the reader…" (Weinberg & Wiesner, 2011, p. 50). The framework employed focuses on text-reader transaction and the features of the text and reader that shape this transaction. It takes into consideration the *intended, implied,* and *empirical* reader. In other words, the author's intended audience, the audience that would truly understand the text, and the actual audience, respectively. When the three readers do not match, or even when just the implied and empirical readers do not correspond, the success of the book in terms of student comprehension and engagement is lessened. The implied reader can be identified by expert readers, in this case knowledgeable mathematicians, as they determine what behaviors and capabilities are required by the empirical reader to make the text effective (Weinberg & Wiesner, 2011).

1.2 Why study example spaces?

Another aspect of textbooks that can influence reader understanding and interaction, and which also can illustrate the intended, implied, and empirical reader, is that of example spaces. The creation of examples is essential in the teaching and learning of mathematics. They are used for reference and as a means to generate other examples, conjectures, and perceptions. Examples clarify and provide context and reference (Bills & Watson 2008; Alcock & Inglis, 2008; Michener, 1978). Bills and Watson (2008, p. 77) claimed that "any theory of learning which does not deal with how learners and teachers act with, and on, examples is likely to be incomplete as far as mathematics is concerned." Examples, and non-examples, of a theorem can aid the process of proving the theorem and understanding the conditions involved. Counterexamples give clear and convincing reasons against a hypothesis and can lead to revised definitions or theorems. Example spaces are similarly needed for definitions because they demonstrate the importance and use of particular aspects of the definition.

There are several conditions that need to be addressed in order for an example space to be ideal. To achieve clarity, the examples should differ along a narrow set of parameters.

Variation helps distinguish essential from incidental features. (Fukawa-Connelly, Newton, & Shrey, 2011; Goldberg & Mason, 2008). Interestingly, the knowledge gleaned from being presented with examples does not seem to be as great as when students generate examples on their own (Dahlberg & Housman, 1997). Zazkis and Leiken emphasize the importance of learner-generated examples, both to the students and the instructor who is trying to evaluate student comprehension. Through example generation and discussion by prospective secondary mathematics teachers, the authors describe how the research subjects view mathematical definitions and the components of a correct definition (2008). Watson and Shipman underline the importance of learner-generated examples; "if students generate examples, reflection on those examples could, through perceiving the effects of the variations they have made, lead to awareness of underlying mathematical structure" (2008, p. 98). Textbooks obviously present the reader with examples, but are the example spaces appropriate? Do the texts include essentially the same examples, leading to a *conventional example space* that teachers then expect their students to become familiar with (Watson & Mason, 2005)? The reader should be given a range of illuminating examples, but also should be led to generate personal examples through the text or exercises. The combination of the two ways to enhance an example space seems to be the best way to increase initial understanding of a concept.

Methods

2.1 Groundwork of the study.

In this study, over a dozen abstract algebra textbooks were considered, some of which were different editions of the same text. The years of first-edition publication ranged from the 1940s to 2010, allowing for a wide range of ideas on what topics are paramount and how best to present them. This article will focus on a subset of the collection which illustrate overall trends. These include the following popular texts: Fraleigh's *A first course in abstract algebra* (1976, 2003), Gallian's *Contemporary abstract algebra* (1994, 2010), Herstein's *Abstract algebra* (1986) and *Topics in algebra* (1964), and the classic textbook, *A survey of modern algebra,* by Birkhoff and Mac Lane (1965). The different authors include, as would be expected, some unique or different topics in their texts. For consistency, this study primarily looked at specific content areas which could be found in all the textbooks, like rings and groups. Prefaces were also useful in determining the authors' intentions. Because the first few chapters set the tone for the entire book, in terms of style and difficulty level, those initial topics will be prominent in this study.

2.2 Employment of reader-oriented theory.

Within the framework of reader-oriented theory, my focus remained on identifying the intended and implied readers and determining whether the author instigates interaction between the text and reader. The empirical reader will not play a large part in this article. While I know anecdotal and personal evidence of the popularity of many of these texts, I have no quantitative data of how many students have used these textbooks, or what type of students they were or are. Many times information given in the preface of the book served as an indicator of the intended reader. Other factors under scrutiny were the language used by the author, the example spaces, the style of proof, and the amount of detail given in explanations. Even within a single book, the level of explanation could significantly differ from one topic or definition to the next, leading to a confusing message of who the implied reader might be. Often, these differences in the expected or necessary abilities of the reader indicate a discrepancy between the implied reader and the intended reader, which could lead to a limited level of understanding by the student. They may not know how to approach the book when some topics and definitions are laid out clearly while others are not. The style of proof can be revealing as well. Differences such as paragraph style versus list style, or more details versus fewer, give evidence of what knowledge the reader is expected or needs to possess in order to comprehend the proof.

The main problem with a lack of consistency between intended reader and implied reader occurred because of varying levels of difficulty within a text, so methods of evaluating intended/implied readers emphasize these discrepancies. I will use the term “high-level” to indicate situations where the material is explained with little detail or when an exercise is quite challenging for the intended level of the book. The term “low-level” will be used when a topic is explained in great depth or with many examples, or there is detailed material on subjects which the intended reader would have likely seen in previous courses.

2.3 Merging reader-oriented theory with example spaces.

A large part of determining the difficulty level of the material came from consideration of the example spaces. For instance, including many examples for one definition, especially when compared to other books, would indicate a lower-level of student proficiency is expected. This type of presentation in a book also deviates from the idea that an optimal example space involves learner-generated examples. There is no need for a student to come up with their own examples when there are so many at their disposal, which means that they may not gain as much depth of understanding. The interaction between the reader and the text could be hindered by qualities like this, as the text is viewed simply as a source of information to be extracted.

The types of examples used can also illustrate how much background knowledge the reader is expected to have. Examples involving matrices or complex numbers, for instance, may assume that the student has seen those topics in a previous class (unless they are introduced within the text itself). Even within the same type of example, different difficulty levels could be present when comparing textbooks. Also taken into account were the examples that were given as exercises. Some authors would ask specifically for the reader to generate an example fulfilling certain conditions, which aligns with the qualities of an ideal example space. Other times the exercises would include dozens upon dozens of problems, providing a plethora of ready examples. Many authors include varying levels of difficulty in their exercises, but some had very challenging problems on par with graduate-level work. A final question when examining the types of examples was whether it matched the intended reader in terms of application problems. For instance, Gallian was trying to connect algebra to the applied sciences like physics or chemistry. His intended reader would be someone who is interested in or needs to know these applications. For the implied reader to match, he would need to actually include application examples, which he does (1994, 2010).

Analysis

Since I am most interested in whether the intended and implied readers match for each book individually, the following results will be organized by author. Then summary remarks will describe the overall conclusions and any interesting connections or differences across authors.

3.1 Gallian

The preface of this textbook states Gallian’s goal as presenting algebra as a contemporary subject, or in other words applying the abstract material to something concrete and useful. He also wants students to enjoy reading the book (2010, p. xi). This indicates a desire for text-reader interaction instead of the student simply extracting information without developing a connection with the material. Gallian includes song quotes, biographies, photographs, and many “real-world” examples to achieve his goal. His intended reader, gleaned from the preface and the book as a whole, is a student being first exposed to abstract algebra. Through reading the book, he wants students "to be able to do computations and to write proofs" (2010, p. xi). This statement implies that the student does not have to be proficient at proofs or abstract computations at the start.

The implied abilities, or the aptitude actually necessary, of the reader in terms of proofs tends to be somewhat mixed. In the first chapter of the book he defines and introduces induction proofs, indicating that the reader doesn't have to be familiar with all types of proofs (Gallian, 2010, p. 13-14). On the other hand, the paragraph style proofs of the text sometimes omit clarifying details. In the first proof of the book, of the division algorithm, a set $S=\{a-bk \mid k$ is an integer and $a-bk\geq 0\}$ is considered without any motivating remarks. To students, this may not be an obvious step to take. Later in the same proof, we are trying to show that $r<b$. To do this, we are told to assume $r\geq b$, which of course is starting a proof by contradiction. This explanatory piece of information, obvious to a mathematician, is not laid out for the reader (Gallian, 2010, p. 4).

In terms of abstract computations, Gallian defines binary operations using multiplication notation from the start, instead of introducing this new idea with "*" or some similar symbol to differentiate it from always being multiplication (2010, p. 40). This can make it harder for students to capture that there are operations other than the usual multiplication under consideration, and that proofs using the multiplication notation would hold true for any binary operation. There is an interesting mixture of explanatory detail, even in the first few chapters of the text.

Both the definition of an induction proof and the division algorithm proof which includes contrapositive arguments, as well as topics like the dot product and $i=\sqrt{-1}$, are in Chapter 0: Preliminaries. The fact that a chapter 0 on preliminaries is in the book indicates that a lower-level of student could be a reader, but the style of proof delivery and lack of introductory notation would indicate that perhaps a higher-level of understanding is needed (Gallian, 2010). Consideration of example spaces illuminates the level of the implied reader as well.

After defining a group, Gallian covers twenty examples, five of which are non-examples. Interestingly, Gallian does stop after example thirteen and tell the reader "with the examples given thus far as a guide, it is wise for the reader to pause here and think of his or her own examples. Study actively! Don't just read along and be spoon-fed by the book." (1994, p. 38, 2010, p. 44). Of course, then seven more examples are immediately given to the reader. Associativity is not proven for any example, but one of the non-examples (the integers under subtraction) is not a group because associativity fails. The reader is asked to prove this as an exercise. There are seven other exercises about determining whether a set and operation form a group, and one requesting the reader generate an example of a group with 105 and then 44 elements (2010, p. 52-55).

In this book, Gallian tries to instigate text-reader interaction and seems to intend that the reader is capable and willing to participate, while at the same time implying that the reader is not or will not by giving numerous ready examples. Although the book is an introductory text to abstract mathematics, clarifying notation and details are omitted. The types of examples, in terms of application problems, often align with the goal stated in the preface of demonstrating algebra's usefulness.

3.2 Fraleigh

As found in the preface, Fraleigh's self-proclaimed goal is essentially to teach as much content as possible. Since not all of the material can be put on a board in class, put it all in a book instead. He expects that readers have studied calculus and probably linear algebra. In order to fit in more algebra, Fraleigh has little in the way of preliminaries, including only sets and relations, which is considered to be review. Because he believes that it will be most students' first exposure to such axiomatic and abstract material, he has "extensive explanations concerning what we are trying to accomplish, how we are trying to do it, and why we choose these methods" (2003, p. vii).

Fraleigh also wrote a preface for students, in which he gives study tips and emphasizes the importance of examples. By the end of this preface, he has defined exactly how he intends for readers to interact with the material: read through once (skipping proofs), gain a real understanding of the statement or theorem, then go back and read again in more detail, and finally try the exercises (2003, p. xi). By the end of the two prefaces, Fraleigh's intended reader is clear. He or she is a student who has had linear algebra, who doesn't require preliminaries but does need in-depth explanations of this material, and who will read through the text multiple times in order to truly understand it.

The first section of chapter one of Fraleigh's book is basically a list of examples involving complex numbers. Although he considers this review, there are eight pages of detailed material in the section, which indicates that the implied reader and intended reader are not quite the same. He expects students to know about complex numbers, but doesn't imply that they do by the actual text. In the next section, binary operations are introduced. Unlike Gallian, Fraleigh does use the "*" notation instead of multiplication, at least until several sections later when subgroups are defined. As with the complex numbers topic, this notation associates with a lower-level of reader, which tends to match with his intended reader who has never seen abstract material like this (2003).

On the other hand, important concepts like *congruence modulo n* and *isomorphic* are given within examples and paragraphs. Modulo *n* is connected to two short examples, isomorphic to a few more, before exercises are given on the concept. Isomorphism is given as a formal definition in a later section (Fraleigh, 2003). This is a high-level approach, especially compared to other texts that will spend a whole chapter on modulo *n* or isomorphic when first introduced (Gilbert & Gilbert, 2009, Hungerford. 1990, Gallian, 2010, Bergen, 2010).

After defining a group, Fraleigh lists nine examples and three non-examples. None of these prove or disprove associativity. There is a nice variation of parameters demonstrated, though. One non-example is the positive integers under addition. It is stated that no identity exists. The next example is the same except includes 0, which means there is an identity. Fraleigh points out that it is still a non-example because there is no inverse for 2. Then the following example is of the entire set of integers under addition, which is, of course, a group. Finally, he changes the parameter of operation by considering the positive integers under multiplication, which is not a group because 3 has no inverse. This type of variation makes clear to the reader what is allowed to change and what sort of effect a change can have, but still leaves many other sets and operations for the student to generate as examples. However, some of those possibilities, such as the rational numbers and matrices, are used by Fraleigh in the rest of the examples. Then there are 16 exercises asking the reader to determine whether a set and operation form a group and one exercise requesting the generation of an abelian group of order 1000. Many of the exercises involve the same sets of numbers, matrices, and operations that were used in the section examples, which makes the homework more accessible to a low-level student. This also means that Fraleigh generally stays within the conventional set of examples used by mathematicians (Fraleigh, 2003, p. 38-49).

Further on in the book, some very high-level concepts appear. He includes chapters on homology and other topics that are often considered graduate-level. Some of the exercises are quite challenging. For instance, after defining rings the exercises include idempotent, nilpotent, Boolean rings, and proving the Chinese Remainder Theorem, all of which are defined within the exercise (Fraleigh, 2003, p. 176-177).

While Fraleigh is often adept at keeping his intended reader and implied reader on the same level, there are some parts of the book that indicate discrepancies. His example spaces follow several of the conditions laid out for an optimal example space, including small variations of parameters and learner-generated examples. The sheer number of examples, however, may deviate from this ideal, while also bringing the implied reader down to a lower-level since they need not be able to come up with those examples on their own.

3.3 Birkhoff & Mac Lane

Considered by many as the inaugural textbook for undergraduate abstract algebra, the first edition of Birkhoff and Mac Lane's text was published in 1941. My edition was the 1965 version, which according to the authors adheres to the same basic philosophy as the original, which is in fact quoted to be "… to express the conceptual background of the various definitions used. We have done this by illustrating each new term by as many familiar examples as possible. This seems especially important in an elementary text, because it serves to emphasize the fact that the abstract concepts all arise from the analysis of concrete situations." (1965, p. v). The authors go on to describe the variety of the exercises, which allow the text to apply to students of the undergraduate or graduate level. Their intended reader covers a fairly large range of possibilities, which they extend to include students directly out of high school (if only some chapters are considered) to those using the book simply as a reference when engaging in research in other subjects, such as physics, chemistry, or engineering.

Due to the range of the intended reader, we would expect to see a range of low-level and high-level material. Low-level certainly appears when the authors spend significant time on ideas like bounds, complex numbers, and linear algebra, which are thus not assumed as part of the knowledge base of the reader. The exercises are less numerous than newer texts, and also exclude complicated proofs and problems (Birkhoff & Mac Lane, 1965). The first proof of the book is actually presented in a two-column style, as shown in figure 1. None of the other textbooks use this style of proof, which many mathematicians would consider low-level. Even this text ceases after the first few proofs and defers to the more common paragraph style. Birkhoff and Mac Lane imply that the reader need have no knowledge of formal proofs, which is perhaps aimed at the intended high school reader. A graduate level student, of course, should be relatively comfortable with formal proofs.

Figure 1

First proof in (Birkhoff & Mac Lane, 1965, p. 3)

RULE 1. $(a+b)c=ac+bc$, for all a,b,c in ring R.

Proof. For all $a,b,$ and c in R:

1. $(a+b)c=c(a+b)$	(commutative law, mult.)
2. $c(a+b)=ca+cb$	(distributive law)
3. $(a+b)c=ca+cb$	(*1,2*, transitive law)
4. $ca=ac,\ cb=bc$	(commutative law, mult.)
5. $ca+cb=ac+bc$	(*4*, uniqueness of addn.)
6. $(a+b)c=ac+bc$	(*3,5*, transitive law)

Birkhoff and Mac Lane introduce rings as the first topic of the book (1965, p. 1). The integers are mentioned as fulfilling the postulates, but no examples are explicitly worked out or described for rings in particular. Instead, an integral domain is quickly defined and the example $\mathbf{Z}[\sqrt{2}]$ given for this definition. While the rest of the first and second sections contain several two-column proofs, there is a noticeable lack of examples (mainly, none other than **Z**[√2]). They conclude with 10 total exercises, three of which ask the reader to determine whether or not they are being presented with an integral domain (1965, p. 1-7).

Groups are introduced through an example of transformations and the mention of the usual rational, real, and complex numbers under multiplication. After the formal definition of an abstract group, there are no other examples in the section, although some exercises ask the reader to determine whether a set and operation form a group. Many of the conventional examples like the rational and irrational numbers under addition, which the other authors describe in the section text, are given as exercises (Birkhoff & Mac Lane, 1965, p. 118-122). Many of the examples in this book defer to the commonly considered sets of numbers like the integers, reals, and complex systems. Thus, Birkhoff and Mac Lane maintained the goal of giving as many familiar examples as possible. Often, the authors follow the rule that allowing the reader to generate examples is more beneficial, as we can see through the low number of examples in the section text. Still, most of the exercises, including the proofs, remain at a low-level of difficulty (Birkhoff & Mac Lane, 1965).

Several aspects of this text point to generational differences in how abstract algebra should be presented and what material should be included. From the thought of defining rings before groups to using a two-column proof, opinions on these methods has changed over time. Presenting fewer examples and exercises, which in general are less difficult than many of the exercises and examples in later texts, points to a lower-level of ability on the part of the reader. Since Birkhoff and Mac Lane believed that this book could be used in a graduate course, it would seem that either the material that is included in an undergraduate versus graduate level class has changed or the authors did not match implied and intended reader.

3.4 Herstein

Another popular abstract algebra textbook is *Abstract Algebra* by I.N. Herstein, which is similar in content to his earlier *Topics in Algebra*. In the preface, Herstein supposed that some readers would be future mathematics researchers, and for them this book would be only an introduction. Other intended readers were those who simply want to keep current with modern mathematics, and perhaps use the ideas in their research with other subjects. Herstein claims that the book is self-contained, except for the last two sections which require some knowledge of complex numbers and calculus (Herstein, 1986, p. vii-viii). Chapter 1 begins with remarks on the fact that Herstein assumes this will be many readers first contact with abstract mathematics. He soon after mentions an expectation that the reader is familiar with set theory (1986, p. 1-3). From these descriptions, we can get a fairly clear picture of the intended reader.

In the first chapter, the implied reader seems to match. Set theory is given but three pages of discussion, and much of the point of the first few homework collections is to get readers more comfortable with the abstraction of concepts and definitions connected to set theory, binary operations, and mappings. An interesting point about Herstein's example spaces are that they do not always fit into the conventional setting of most other texts. He moves beyond the usual integers, complex, matrices, etc… and uses sets of boxes and triangles, states and citizens in America, and even the Kennedys (1986, p. 3-9). Through examples such as this, the book becomes accessible to those who are less experienced mathematically, which matches his intention of it being self-contained. Additionally, the exercises in the text are split into the categories: easier, middle-level, and harder (Herstein, 1986). Hence, the reader is more aware of what is expected of them and where they should focus their attention. A future mathematics researcher could try the harder problems, while someone who is just trying to keep up with the general trends of modern mathematics may only try the easier. Herstein is able to reach the wider audience that he intends through presentations like this.

Moving on to the introduction of groups, Herstein presents the definition using the general "*" notation for an operation, and even further explains it so that there is no confusion about the fact that this is not necessarily multiplication. Twelve examples and three non-examples follow, even listed under those headings for extra clarity. In terms of small variations in the parameters of the examples, Herstein performs this nicely by moving from the integers to the rationals to nonzero rationals to the positive reals, consecutively. Eight exercises ask the reader to prove whether or not a set and operation form a group. The harder exercises also request generated examples (1986, p. 46-55).

Herstein intended this book to be read by a range of students, and although it includes both higher-level and lower-level details and concepts, the way that it is arranged allows the implied reader to correspond to the intended. Furthermore, the example spaces are unique and seem to, in large part, match the components described for an optimal example space.

Summary

The undergraduate abstract algebra textbooks studied in this article were mostly successful in matching intended and implied readers. In some cases, the level of details or difficulty were not perfectly in line with the proposed level of the reader. Fraleigh had graduate-level exercises and topics for an introductory text. Gallian left out clarifying notation and details, necessitating a higher-level of reader than he seemed to intend. Birkhoff and Mac Lane believed their book could be graduate-level, but the examples and exercises generally do not equal the difficulty of the other texts.

Some of the authors clearly realized that reader-response, the interaction between reader and text, is important. Gallian actually interrupts an example space to stop and tell the reader to come up with their own. He also includes the fun quotes and historical facts that he feels makes the book more engaging. Fraleigh outlines how he expects the reader to approach the book, which is not that they simply extract information but instead read through it enough times that they truly understand the concepts. Herstein's unique example space, using popular culture and other real-world topics, connects the abstract material to personal traits of the reader.

As a final comment on the textbooks, I want to emphasize the conventional example space. As mentioned, Herstein is one of the few authors who includes original examples. Gallian does so, as well, through his attempts to give interesting applications of the abstract material. On the whole, though, almost all of the example spaces for the definitions and theorems of the textbooks used the same sets and operations. Of course, the integers, rationals, reals, and complex numbers under multiplication or addition occur frequently. Matrices often come up as an example for the same topics. When an irrational number is needed, the authors tend to choose $\sqrt{2}$, leading to many example spaces including $\mathbf{Z}[\sqrt{2}]$ or $\mathbf{Q}[\sqrt{2}]$. Zazkis and Leikin (2007) also found that this was the irrational number consistently used by mathematicians. Authors often ignore the same characteristics of a definition in their examples, such as the associativity of a group. In all, there does seem to be a definite conventional example space for abstract algebra textbooks.

This study of abstract algebra textbooks is meant to be a beginning step in the direction of research on upper-level mathematics textbooks, especially those for pure mathematics. It needs to be taken further by considering more of the textbook as a whole, instead of focusing on a few sections and chapters. Also, the framework can be employed on other texts and subjects, and the framework itself could be extended to include new methods or important ideas.

References

Alcock, L. & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69(2)*, 111-129.

Bergen, J. (2010). *A concrete approach to abstract algebra*: *From the integers to the insolvability of the Quintic.* Burlington, MA: Elsevier.

Bills, L. & Watson, A. (2008) Editorial introduction. Special Issue: Role and use of exemplification in mathematics education, *Educational Studies in Mathematics*. 69(2): 77-79.

Birkhoff, G. & Mac Lane, S. (1965). *A survey of modern algebra*. New York: Macmillan.

Borasi, R. & Seigel, M. (1990). Reading to learn mathematics: New connections, new questions, new challenges. *For the Learning of Mathematics, 10*(3), 9-16.

Crouch, R. & Walker, E. (1962). *Introduction to modern algebra and analysis*. New York: Holt, Rinehart, and Winston.

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics, 33(3)*, 283–299.

Eco, U. (1970). *The role of the reader: Explorations in the semiotics of texts*. Bloomington, IN: Indiana University Press.

Fraleigh, J.B. (1976). *A first course in abstract algebra*. Reading, MA: Addison-Wesley.

Fraleigh, J.B. (2003). *A first course in abstract algebra*. Boston, MA: Addison-Wesley.

Gallian, J. (1994). *Contemporary abstract algebra*. Lexington, MA: D.C. Heath and Co.

Gallian, J. (2010). *Contemporary abstract algebra*. Belmont, CA: Brooks/Cole.

Gilbert, L. & Gilbert, J. (2009). *Elements of modern algebra*. Belmont, CA: Brooks/Cole.

Goldenberg, P. & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics, 69(2)*, 183–194.

Herstein, I. N. (1986). *Abstract algebra*. New York: Macmillan.

Herstein, I. N. (1964). *Topics in algebra*. New York: Blaisdell.

Hungerford, T. W. (1990). *Abstract algebra: An introduction*. Philadelphia, PA: Saunders College Publishing.

Jacobson, N. (1974). *Basic algebra I*. San Francisco, CA: W. H. Freeman and Co.

K-12 Mathematics Curriculum Center. (2005). *The changing mathematics curriculum: An annotated bibliography*. Newton, MA: Education Development center.

Kang, W., & Kilpatrick, J. (1992). Didactic transposition in mathematics textbooks. *For the Learning of Mathematics, 12*(1), 2-7.

Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational Studies in Mathematics, 52(1)*, 29-55.

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.* (pp189-202) Washington: MAA.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.

Michener, E. (1978). Understanding Understanding Mathematics. *Cognitive Science*, 2(4), 361-383.

Morgan, C. (1996). "The language of mathematics": Towards a critical analysis of mathematics texts. *For the Learning of Mathematics, 16*(3), 2-10.

Fukawa-Connelly, T., Newton, C., & Shrey, M. (2011) Analyzing the teaching of advanced mathematics courses via the enacted example space. In (Eds.) S. Brown, S. Larsen, K.

Marrongelle, and M. Oehrtman, *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education*, Vol. 1, pg 1-15. Portland, Oregon.

Reys, B. J., Reys, R. E., & Chaves, O. (2004). Why mathematics textbooks matter. *Educational Leadership, 61(5*), 61-66.

Robitaille, D. F., & Travers, K. J, (1992). International studies of achievement in mathematics. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 687-709), New York: Macmillan.

Rosenblatt, L. (1938). *Literature as exploration*. New York: Appleton-Century.

Rotman, B. (2006). Toward a semiotics of mathematics. In R. Hersh (Ed.), 18 *unconventional essays on the nature of mathematics* (pp. 97-127). New York: Springer.

Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: learners generating examples*. Mahwah, NJ: Lawrence Erlbaum Associates.

Watson, A. & Shipman, S. (2008). Using learner generated examples to introduce new concepts. *Educational Studies in Mathematics, 69(2)*, 97-109.

Weinberg, A. & Wiesner, E. (2011). Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics, 76*(1), 49-63.

Zazkis, R., & Leikin, R. (2008). Exemplifying definitions: A case of a square. *Educational Studies in Mathematics, 69*(2), 131-148.

AN EVOLVING GRAPHICAL IMAGE OF APPROXIMATION WITH TAYLOR SERIES: A CASE STUDY

Danielle Champney
University of California, Berkeley

Eric Kuo
University of Maryland, College Park

This paper will take a close look at the construction of a graphical image for reasoning with approximation in the context of Taylor series. In particular, it is a comprehensive case study of the genesis and evolution of an image created by one student, who draws extensively on other images and knowledge from calculus and physics to supplement gaps in his understanding of Taylor series and reason with Taylor series approximation tasks. His process resulted in a graphical image that was leveraged to build knowledge and reason with the situation, even while lacking key considerations that are central to an understanding of Taylor series. In this paper, we speak not only to considerations of a student's understanding of this particular content. This work also provides a detailed examination of the processes of constructing a graphical image used for problem solving, for which it was necessary for the student to obtain and utilize evidence to amend that graphical image.

Keywords: Taylor series, graphical representation, calculus

Introduction

Taylor series comprise a large portion of a calculus curriculum, and for good reason – they are a building block that many other disciplines use as a jumping off point for more advanced topics. But, students' reasoning with this topic is vastly understudied in the current literature. What sense do students make of Taylor series? Do they have any image at all for Taylor series and what they're used for? The little research on students' understanding of Taylor series speaks mostly to broad themes of characterizing expert/novice strategies (e.g. Martin, 2009), dispositions toward reasoning with them (e.g. Alcock & Simpson, 2004 and 2005), grappling with formal definitions of convergence (e.g. Martin et al, 2011), or use of technology in instruction (e.g. Yerushalmy & Schwartz, 1999; Soto-Johnson, 1998). That is, most of these important studies on students' use and understanding of Taylor series take a global perspective, examining general themes, post hoc[1]. But to develop a robust knowledge of the concept of Taylor series requires the synthesis of many previous calculus content topics, woven together and used appropriately, to form a more complete image. The question of *how* students synthesize their prior knowledge and arrive at *their* image has not been studied. That is, we have little idea about how students *construct* an understanding of Taylor series from less formal prior calculus notions, and how they attribute meaning to particular aspects of whatever representation of a Taylor series they espouse.

Visual images of Taylor series

[1] One exception, Martin et. al (2011), does focus on students' construction of ideas of pointwise convergence, but with a focus on formal definitions.

Though it is not always students' tendency to produce visual images for Taylor series tasks, many do (including the student discussed here). Access to students' visual images, supplemented by their descriptions and explanations, can provide additional insight into how they are constructing an understanding of topics such as Taylor series, as visualization is "a fundamental aspect to understanding students' constructions of mathematical concepts" (Habre, 2009). That is, students' graphs, pictures, and other inscriptions can provide a physical referent of what they are understanding of the concept. With topics such as Taylor series, for which students often use mathematical and academic language imprecisely in a genuine attempt to find words to characterize what they understand of the topic (see Monaghan, 1991 for examples in the context of limit), these visualizations become even more important to study. Martin (2009) showed unsurprisingly that mathematicians were more fluent than novices in using graphical representations, both in their construction and interpretation, in the context of Taylor series. His work makes clear that "many students do not have a good visual image, if they have any visual image at all, of the convergence of Taylor series" (p. 288). Biza, Nardi, and Gonzales-Martín (2008) agree, citing an additional lack of useful imagery in textbooks chapters that students may use for reference. In our experience and works in progress, which align with Alcock and Simpson (2004 and 2005), many students do in fact turn to visual images to explain and reason with Taylor series tasks. In fact, Alcock and Simpson (2005) also demonstrated that even "non-visualizers" may have a reliable graphical image, but tend to not call on it. So, in this paper, we endeavor to study, with a moment-to-moment analysis, the creation of one student's graphical image that *he* chose to use to play out his reasoning with Taylor series approximation tasks.

Theoretical Framework

In this paper, we investigate a student's reasoning around Taylor series through what we refer to as "graphical images." We refer to a "graphical image" as a particular type of visual image, which is made of any set of inscriptions that can include multiple formal mathematical graphs and informal diagrams and pictures overlapping on the same coordinate axes. In using the term "image," we do not mean to align with other uses of the term in the literature (e.g. Tall & Vinner's (1981) "concept image"). We more closely align with a sense of "graphical and pictorial representations," though simply using the term "graph" does not, to us, capture all of the inscriptions that one may make when explaining something graphically. What we share with the representations literature is the interest to "investigate the potential of students' self-generation and elaboration of visual representations, to enhance and advance their understanding of difficult conceptual domains" (Parnafes 2009, p. 147) and "focus on the spontaneous generation and use of domain-specific diagrams during reasoning," (Kindfield 1993, p.1).

In order to investigate the graphical images of one student, we study not just the final products of his reasoning, but rather the student's reasoning *as he constructs* his explanations. Even for students who have demonstrated success when assessed on certain topics in their math classes, reasoning around these topics may not necessarily be clean and consistent. This is often true with regard to the topic of Taylor series, for which standard assessment items commonly require little more than a well-rehearsed procedure for arriving at a solution. Habre (2009) discovered that even multiple exposures to the topic of Taylor series, at varying levels of mathematical sophistication, are often insufficient for even a broad comprehension of the material. Thus, knowing how students build their understanding can put into perspective some of the issues that persist around this topic. Our goal is to understand the conceptual dynamics that occur in the moment-to-moment reasoning in an interview setting (Sherin, Krakowski & Lee, 2012).

This approach has two noteworthy affordances. First, this allows us to understand pieces of a student's reasoning in relation to the whole. What can look like a firmly held misconception at one moment in time can merely be a temporary stepping stone in a student's reasoning as it develops. Second, it can give us access to not just the final products of student reasoning, but also *the process* through which one student's graphical images are developed and refined. More specifically, how additional information gathered during reasoning and problem solving, as well as a student's prior knowledge, work together to guide the development of those graphical images. Although the details of such a story can't be generalized beyond the student in this case study, it can elucidate certain general mechanisms driving student's moment-to-moment reasoning that can suggest broader investigation.

Although this methodological approach is present in the literature (e.g. Schoenfeld, Smith, & Arcavi, 1993), our study addresses a dearth in the literature as few studies in the mathematics education literature on the topic of Taylor series have investigated student reasoning in such a fine-grained manner (exceptions include Martin et al., 2011). Additionally, those that have were focused on students' reasoning about formal definitions, with much attention to mathematical rigor. Our purpose for this type of analysis is different in that we are interested in a student's reasoning independent of canonical correctness. In fact, we are especially interested in how students may develop mathematically incorrect explanations into canonically correct ones by incorporating both information gained during problem solving and their prior mathematical knowledge. Furthermore, teaching experiments like Martin et al. (2011) span multiple hour-long sessions and incorporate scaffolding through gradual construction of understanding through a series of designed activities and researcher guidance. In contrast, our study presents one portion of one student's work in an hour-long interview, with little scaffolding beyond the written task prompts and occasional interview requests for clarification or confirmation.

Research Questions

The strength of the case to be presented in this report is two-fold. As it is a detailed examination of the development of a student's reasoning, as it plays out *graphically*, following this student's process with a moment-to-moment analysis can allow for an examination of what he takes as calculus-based and physical evidence for claims he is making in his reasoning, and how those claims are manifested in his graphical image. Second, and much more content-specific, Taylor series literature largely examines students' (graphical, and other) reasoning or presentation at the *completion* of a problem, rather than as it is being *built,* negotiated from one moment to the next. Therefore, the exploration of this case will speak to the following:

(1) How is additional evidence germane to a problem gathered and used to amend a graphical image that serves to represent a particular concept for a student?

(2) In what ways are prior calculus concepts negotiated to construct and attribute meaning to this representation of Taylor series?

With the case presented here, these questions can *only* be addressed for one particular student, but can be used as a model both for future analyses, and to highlight ways in which calculus-based reasoning *can* (and does) influence students' understanding of Taylor series.

Methods and Data Collection

The study makes use of a particular 1.5-hour semi-structured interview with sophomore physics major Joe, who was participating in a larger, related study. Though it will not be discussed in this paper, the purpose of the larger study (Champney, Kuo & Little, in preparation) was to investigate students' consistencies (and inconsistencies) in reasoning around a set of approximation tasks in calculus and physics contexts. The tasks discussed in this paper are the only two in the larger task set that elicit students' thinking about approximation in the context of Taylor series (the text of which appear in Figures 1 and 2). They are intentionally vague, and written to allow participants great freedom in what they choose to attend to as they respond. The interview was videotaped and transcribed for analysis. At the time of the interview, Joe had taken three semesters of calculus and two semesters of physics, and earned grades of "A" in all of them. He was identified by instructors as very competent in the subject matter. Upon completion of data collection for the larger study, Joe's interview stood out for several reasons, of interest here are those related to his construction of Taylor series images. An in-depth, microgenetic analysis of this interaction between the student and the calculus content at his disposal seemed a promising way to examine *change* in his notions of Taylor series approximation on a more fine-grained level than would be ascertained in other assessment situations (Calais, 2008).

The Taylor series about x=0 for arctan(x) is given by: $\arctan(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \ldots$

How big a value can x be, before stopping after the second term is a bad approximation?

Figure 1: Task 1

You have a pendulum made of a metal ball on a string. The string is 1 meter long and the metal ball has a mass of 1 kg. You might know that the approximation for the period of a pendulum for small oscillations is

$T = 2\pi\sqrt{\frac{l}{g}}$ where T is the period of the pendulum, l is the length of the pendulum, and g is acceleration due to gravity (9.81 m/s^2). This equation only holds for small angle oscillations of the pendulum. For larger angles, the period of a pendulum can be found with the following equation:

$$T = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{1}{16}\theta_0^2 + \frac{11}{3072}\theta_0^4 + \ldots\right)$$

where θ_0 is the angle of displacement of the pendulum from vertical in radians. You want to calculate the period of oscillation for this pendulum. How big can the angle of displacement of the pendulum be before the equation for small oscillations isn't a good approximation of the period?

Figure 2: Task 2

A complete narrative of Joe's work with Task 1 will highlight the nature of the transition points in his thinking, as they are played out on his evolving graphical image. This analysis, then, should illuminate both how Joe uses his additional evidence to refine his image, and how that image represents the meaning of approximation with Taylor series (according to him). To carry out an analysis of this entire interview, it was broken into episodes during which Joe is appealing to a stable version of his graphical image. Within each episode, it is then instructive to trace his thinking and evidence for his claims, both as he discusses them and as he amends his image based on those claims. When he abandons one image for another structurally different version, a new episode begins.

Results and Analysis.

In this section, we present such a narrative of Joe's interview, describing each stable graphical image in detail, including the relevant discussion that shaped it, and the conversation around it that was used for clarification. Each stable graphical image will be referred to, in turn, as Graphs 1, 2, and 3, each for Task 1. After describing Joe's work on Task 1, relative to our methodological aims, we turn to a discussion of his work on Task 2, and the ways in which his final image (Graph 3) provided a stable starting point for reasoning with Task 2.

Task 1 Preliminaries. Upon first inspecting Task 1 (Figure 1), Joe identifies several facts that he knows about arctangent – that arctan(1)=π/4, the graph of arctangent is concave down, as x gets bigger, arctan(x) changes by smaller amounts, and arctangent has an asymptote at y=π/2 – which allow him to construct an accurate graph of that function. More importantly, Joe identifies several key components about his understanding of Taylor series:

> *Joe: As I understand with Taylor Series the first terms are essentially the biggest terms and the terms after that tend to take into account smaller deviations. For example, when you derive physical formulas of Taylor series you'll typically stop after the second or third term because ... say, 1 over 129 x to the [power of] 129 aren't going to make that big of a difference.*

After setting down his background knowledge about the graph of arctangent, and terms of a Taylor series, Joe makes his first move toward addressing the content of Task 1. He chooses to plug x=1,000 into $y = x - \frac{1}{3}x^3$, and compare with what he knows to be the approximate value of arctangent at that same x-value. While it is unclear why Joe chose the particular value of 1,000, his intention to pick a sample value and examine the differences in the two functions for that particular value is established early.

Background to the Graph 1. The information that Joe gleans from plugging in x=1,000 provides a backdrop for the first graphical image that he produces.

> *Joe: If you take x = 1,000 the first two terms will give you 1,000 minus ... it gives you, well... It gives you a very large negative number. Roughly, -3.3 times ten to the eighth, I believe. That is obviously very far from the answer.*
> *Interviewer: Which should be what?*
> *J: Um, roughly pi over 2*

Noting that $y = x - \frac{1}{3}x^3$ gives a very large negative number for x=1,000, while arctangent should produce a value close to π/2, Joe makes a decision to try to focus on when/where the graph of $y = x - \frac{1}{3}x^3$ is sufficiently near to π/2. In order to find the values of x for which this is reasonable, Joe decides to use his existing (correct) graph of arctangent to demonstrate how such a range of x values could exist.

Graph 1. In order to make sense of his ideas that $y = x - \frac{1}{3}x^3$ should at some point be near $\pi/2$, Joe constructs what he will from this point forward refer to as an "interval of confidence." Note that Joe does not mean this in the normative, statistical sense of a confidence interval[2]. It is unclear if he even has a sense of what this term traditionally means, as he expressed that he had never taken a statistics course. Rather, Joe is using this notion to describe a sort of tolerance band around the desired values of a function, which would define 'where' an approximation is appropriate or not. In his words,

> *J: you, essentially you would … need to define some interval of confidence. Like say you want to be within 0.1 of pi over 2. So pi over 2 is ... about 1.57. So say you want to be between 1.47 and 1.67.*
> *I: Okay.*
> *J: So then you would have to try to find a value of arctangent of x. Well the thing is if you take ... if x=0 its obviously not at pi/2.*
> *I: Should it be?*
> *J: Um, well no. I'm just thinking, since it starts outside of the range, you'll want to see when it first enters the interval of confidence and you'll want to see where it exits.*
> *I: Oh I see. Okay, so you're looking for some kind of tolerance around some x values where is crosses this?*
> *J: Yeah, essentially there should be some range of x values where the Taylor approximation with the first two terms will give you an answer between 1.47 and 1.67.*

This dialogue prompts Joe to draw dotted horizontal lines around his asymptote on the graph of arctangent, above and below, demonstrating that a reasonable approximation of the function will produce values that are within that band (see Figure 3). Further explaining, Joe proceeds to draw his version of what the approximation might look like (the thick curve on Figure 3), stating

> *J: And um, and then once it enters the interval of confidence you begin to encounter the possibility of it being a bad approximation, so then once it leaves that interval of confidence you know that it's become a bad approximation …*[transcript snipped]*… So to express it on the graph, that band of f(x) values would be like that, and then … the f(x) values you get using the first two Taylor approximations will probably look something like this I'm thinking. I'm not sure.*

As he talks through it, Joe points on his graph to the places where his approximation crosses into and out of his confidence interval, as the points where the approximation first "encounters the possibility of being bad" and then has "become bad."

---

[2] And therefore, as we refer to "confidence interval" or "interval of confidence" in the remainder of this paper, we adopt Joe's meaning.

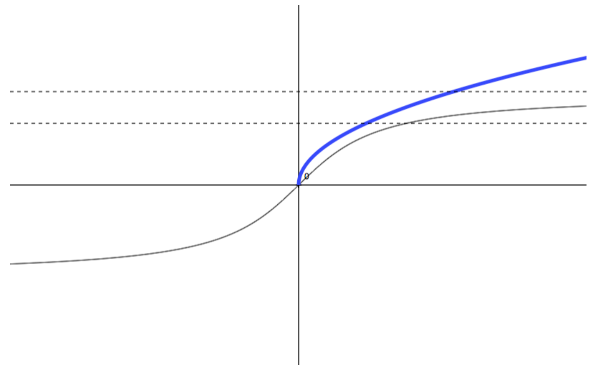


Figure 3: Graph 1, confidence interval around pi halves

Though he explains in great detail why he believes this is a good strategy for determining when the approximation (a cubic) would represent a reasonable approximation for arctangent, and shows great skill in graphing arctangent and reasoning about end behavior, ranges of values, and other graphical notions, it is clear that from this first image that Joe has neither recognized several key considerations about Taylor series, nor produced the correct graph of $y = x - \frac{1}{3}x^3$.

If one simply stopped engaging with Joe at this point, taking his work as evidence of his understanding of Taylor series, the conclusion would likely reflect his misunderstanding of the meaning of the center of a Taylor series, and also that his more basic graphing skills for simple polynomials was lacking. While we may see some semblance of a good idea in his decision to bound the function, Joe has placed the bound in an inappropriate place that has nothing to do with the question being asked. Though a perfectly reasonable question (though futile in this task) may have prompted students to think about how many terms would be necessary to approximate end behavior of a function with its approximation, this is simply not the question asked during Joe's interview. However, Joe's initial graph (Graph 1), for him, proved only to be a stepping-stone toward a more successful understanding of the topic. The flawed visualization that he created allowed him to grapple with inadequacies that *he* recognized, in his graphical image, in order to amend how he was thinking about the topic.

Transition to Graph 2. Rather than intervene to correct aspects of Joe's graphical image, the interviewer remains silent and permits Joe to grapple with it. And, upon further reflection, Joe recognizes two problems with this representation. First, he recognizes that it "starts outside the range" - that is, he notices that the point that the two graphs share (the origin) is outside of his interval of confidence. He chooses to explain this away and not act on it, not recognizing the importance of the 'center' at x=0. That is, this discrepancy is explained away by assuming that you have a good approximation up until you enter the interval of confidence, and not based on reasoning about the Taylor series itself.

However, Joe *does* act on a second problem – having previously plugged the value of x=1,000 into the first two terms of the approximation, Joe expresses dissatisfaction with his graph of the approximation, stating that he thinks "[the cubic] will go off to negative infinity." It is at this point that Joe shifts his thinking, uses a calculator to graph $y = x - \frac{1}{3}x^3$

Graph 2. With a correct graph of $y = x - \frac{1}{3}x^3$ at hand, Joe recognizes that the approximation will never even reach his interval of confidence. Amending his image to Graph 2 (Figure 4), Joe's attention shifts to more local features of the graph. Rather than choose extreme values like x=1,000 for comparison, or focus on end behavior of the two graphs, Joe considers features of the graphs like the maximum of the cubic function. Noting that "arctangent is strictly increasing," and that the cubic has a maximum, Joe posits

> *J: actually I was thinking that might make it much easier because for positive values of x arctangent of x is strictly increasing. But um, well, the cubic* [approximation] *is decreasing after a certain point, so once it passes that point you know it is rapidly becoming a bad approximation.*

While he had originally convinced himself thoroughly that tolerance bands around $\pi/2$ were appropriate, new evidence (both numerical and graphical) prompted Joe to, for the moment, abandon the idea of a confidence interval. No longer concerned with the asymptote, his focus shifts to the increasing/decreasing features of the two functions in question, which are graphed accurately on his Graph 2. That is, he gathered evidence that caused an amendment in his graphical image, momentarily foregoing end behavior to accommodate what he knows about a more local feature of the graph.

Returning to his previous idea of comparing the two graphs at a point, Joe now uses his new graph to compare the graphs at x=1, the local maximum of $y = x - \frac{1}{3}x^3$, concluding

> *J: Yeah, so essentially when x=1 arctan of x equals pi over 4 which is roughly 0.78, And then um let's say, let's use capital A as approximation of x. That is two thirds, which roughly equals 0.67. So by the interval of confidence I suggested, 0.1 on each side, it's already a bad approximation.*

With this calculation, the very notion of interval of confidence, for Joe, has shifted. What previously referred to a horizontal band around infinite behavior has now been localized, to a criterion to which comparison of individual points should be subjected. With this new frame, and new graph to support it, Joe's attention is on comparison of individual points according to a 0.1-tolerance, using the local maximum of $y = x - \frac{1}{3}x^3$ as a first point of comparison.

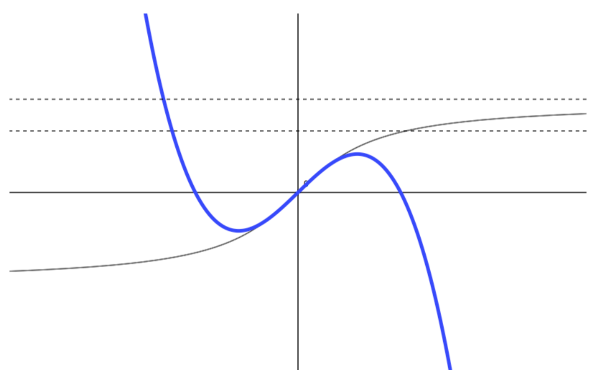


Figure 4: Graph 2, with accurate graph of the approximation

The idea of confidence interval, for Joe, was adapted in light of the new evidence of the correct graph of the approximation function. When the approximation was accurately graphed onto Graph 2, the 0.1-confidence interval turned into a 0.1-criterion for comparison at a point that appeared salient. This point was, for Joe, a logical starting value, in light of his prior calculus knowledge about increasing and decreasing functions ('it must be a bad approximation if the functions are behaving entirely differently'), and the graphical evidence in front of him.

Background to Graph 3. The focus on point-wise comparison marks another shift in Joe's thinking. While his first idea of producing an interval of confidence around the end behavior was incorrect, the notion of producing what Joe refers to as an interval of confidence was not a particularly bad strategy. Further, the accurate graph of $y = x - \frac{1}{3}x^3$, which prompted Joe toward comparison at particular *x*-values, allows the two ideas to be merged into what Joe will construct as his third and final graphical representation for the task.

Graph 3. With his new evidence, Joe's final strategy is to use an adaptive version of his interval of confidence, to allow for the pointwise comparison between the functions at any point. That is,

> *J: I suppose you would want ... You could use a different ... I think that a better idea would be to use a different ... Try to use a different ... Try to figure out what it will be at each point along the way. So then you can adjust your interval of confidence. So instead of like* [the previous bounds around the asymptote]*, it will look more like* [Figure 5].
> *I: Oh I see, okay … So you are just sort of [drawing]a dotted line around arctan?*
> *J: Yeah. And then you can find where this [approximation] leaves this dotted line.*
> *I: So, let me see if I can understand then. If you look at the green graph you drew, somewhere around here?*
> *J: Yes.*
> *I: What are you saying then about the approximation around there?*
> *J: I'm saying that the approximation becomes some distance away from the actual function that's greater than you want it to be.*

While discussing this idea, Joe draws a dotted "confidence interval" around the entire graph of arctangent, illustrating that this new adaptation of his idea will allow one to examine the distance between the approximation and the actual function for any value of *x*. He has thus arrived at a more normative way of approximating with Taylor series, as exemplified on his final graph.

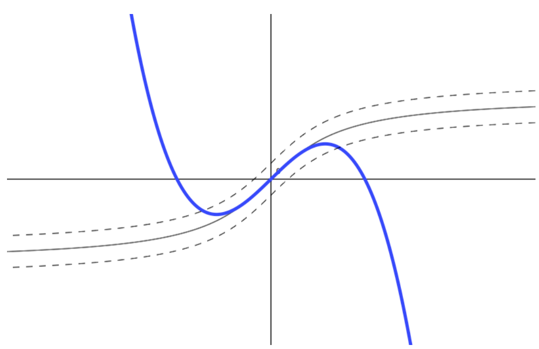


Figure 5: Graph 3, with confidence interval around the graph of arctangent

While in the context of teaching a calculus course, instructors may not use the same sort of graphical image, Graph 3 represents for Joe the way of reconciling his idea from Graph 1 about producing a bound to 'capture' the approximation with the notion from Graph 2 that comparison at particular points is a useful strategy for determining if an approximation is "good" at those points. Thus, Graph 3 represents Joe's final word on Task 1 – a way of thinking about the approximation being "some distance away from the actual function that's greater than you want it to be."

If we were only to examine Joe's Graph 3, much of his process for arriving at that final image would have been masked. If one is only interested in whether or not students can work productively with this particular topic, at the end of the day, then that sort of analysis may be adequate. However, looking at the entire process of Joe's reasoning with Task 1 provided more information than Graph 3 could alone provide. Examining the intermediate images that Joe used to seek, gather, and synthesize information he found relevant to the context of Taylor series, we see the path through previous and current mathematical topics that brought him to his final graphical image.

There is additional evidence (later in the interview) that Joe's final graph for Task 1, while certainly more correct and appropriate than his first attempts, was perhaps the consequence not only of coming to a more normative way of using/viewing Taylor series, but also of attempting to rectify a problem stemming from not understanding the role of the center of a Taylor series. That is, Joe repeatedly identified the idea that the point (0,0) not being within his initial interval of confidence around arctangent was "problematic" and "undesirable," rather than a result of a problem with the interval of confidence, as it is employed in Graph 1. That it should be the starting point from which one would compare the functions seemed to escape him. One potential reason for his choice to put a confidence interval around arctangent, then, is to capture the elusive point (0,0) that he knows the two functions share. As he says at the end of the interview, when reflecting on the entire set of tasks he worked on,

> *J: One thing that struck me as interesting is that for the first problem the starting point is actually rather undesirable. Or, at least that was my first impression of it because you have arctangent of x, which approaches pi over 2. And um, but it starts out at zero, as so as you saw my initial idea with the interval of confidence was to try to find out where it left an interval of confidence around pi over 2. But then, as it turns out it never would actually reach that interval of confidence. So that makes it rather tricky.*

Whether or not this is part of the reason for Joe's amendment to Graph 3, the notion that he wishes to examine the distance between the approximation and the actual function is strong, and it is clear that in so doing, one would want to include all portions of the graph of $y = x - \frac{1}{3}x^3$ that lie within the allowed distance from the graph of arctangent. Thus, his possible inattention or misunderstanding of the idea of a center of a Taylor series is secondary to the more robust way that he has chosen to define and depict (graphically) what he means about "good approximation."

Considering the collection of three graphs that Joe produced when making sense of this approximation with Taylor series task over a period of 20 minutes, it is then possible to trace the particular pieces of mathematical evidence that caused shifts in Joe's graphical image, from a wholly incorrect starting point, to a final and more normative image for how one might successfully explain what it means to use a Taylor approximation, graphically. Navigating numerical, graphical, technological, and contextual evidence, Joe was able to amend his graphical image to reason successfully about Task 1. Additionally, in Joe's work on Task 2, there is evidence that his ideas persisted, based on the way that he reasoned about Taylor series in an entirely different context.

Task 2 Preliminaries. From the very beginning of his work on Task 2 (Figure 2), Joe immediately adopts the stance that is evident in his final work on Task 1 – that is, he endeavors to look at the distance between the actual function and the approximation. Though the constituent parts of the approximation are presented differently in Task 2, Joe immediately recognizes the information as "a Taylor approximation," and proceeds (spontaneously) to graph both the constant function representing the small angle approximation of period, and a graph with positive end behavior to represent "the series."

Joe again shows great facility with graphing functions and relating the functions in the problem. It is questionable what exactly he graphed when drawing "the series," and when prompted he indicated that he drew something that "grows like a quadratic," meaning something with positive infinite end behavior (See Figure 6). After he spends a short amount of time trying to discover some closed form representation of the expanded series, to no avail, Joe decides to try and relate his strategy from the first problem to the one at hand.

Task 2 Graph. Joe's graph for Task 2 underwent no revision throughout the 20 minutes of work with it. After tinkering with the terms of the series in an attempt to "figure out the nature of the series," Joe states

> *Joe: Back to the problem itself... So sort of like we did the first time, you would ... want to have an interval of confidence around the period of oscillation. Though I'm thinking it would be easier to have an interval of confidence around the approximation, since it's constant.*
> *Interviewer: Okay*
> *J: And then you would look to see where the period ... the pendulum left this interval of confidence, and that would be where it becomes a bad approximation.*
> *I: So why would that be where it becomes a bad approximation? Just so I know.*
> *J: Well because ...because if it is a good approximation, then that means the actual period is close enough to 2 pi square root of l over g for ... uh what are your needs.*

That is, producing the graph found in Figure 6, Joe returns to his idea about examining the distance or difference between the approximation and the actual function as a way to measure 'goodness' of approximation, and chooses to illustrate this graphically, by drawing both the constant function, a graph representing the eventual behavior (according to him) of the series, and another interval of confidence to capture the distance between the functions being greater than desired.

In an interesting adaptation, Joe further determines that it is appropriate to associate the interval of confidence around the constant function, rather than the more complicated one, stating

> *J: I'm thinking it doesn't actually matter whether you associate the interval of confidence with the actual function or with its approximation. Since either way, you're looking at how far ... you're looking at the distance between them at any given value of theta, in this case. So since the approximation is a constant function ... I'm thinking we should associate it ... we should apply the interval of confidence to the constant function. So then you have, essentially, the period will be strictly increasing ...*

In this way, Joe has taken his idea of confidence interval – which was developed through his work on Task 1, first as a bound on end behavior, to a localized bound for comparison at various points, to a continuous bound that allows for pointwise comparison across the entire function – and flexibly applied it to a new context, adapting its use, while remaining true to his concept of what it means to him to approximate with a Taylor series. That is, the interpretation of the approximation and actual function being a specified distance away from one another is maintained. And his conclusion is consistent with the final state of his work on the first task, though reversing the roles of which of the approximation and the actual function "cross the interval of confidence," since the function being bounded in Task 2 has switched. That is,

> *J: When the period passes the upper line you know that it has become bad.*

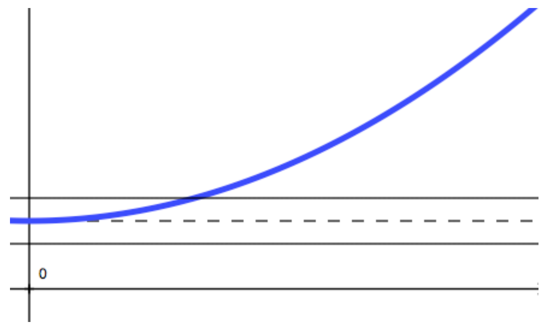


Figure 6: Graph for Task 2

Joe continues to verify for himself that his representation of Task 2 is indeed fitting with his ideas of good and bad approximations, using numerical reasoning (testing various angle measures), graphical reasoning about what would happen for very large theta, and contextual reasoning with the pendulum as a real-world example. However, none of the additional reasoning that he does (more than fifteen minutes of exploratory, unguided probing of the context) prompts any revisions or additions to Joe's graphical image of the task.

Reflection on the idea of interval of confidence. Upon completion of the two tasks, Joe spontaneously begins to reflect on the similarities and differences he noticed. The interviewer asks for further clarification on the notion of using intervals of confidence in the two tasks, to which Joe responds

> *J: I think it's actually easier to see it in the third problem than the first because in the third problem the approximation is constant. Which means ... So the confidence bounds just allow you to think of it as what T, the period, exceeds a certain value. Whereas in the first problem, the function and the approximation... Well the function is an inverse trig function and the approximation is a negative cubic. So it's rather difficult to deal with the error bounds there.*

Joe's use of a confidence interval, then, is consistent with his notion of comparing the two functions in a continuous manner, and flexible enough to adapt to the difficulty of the functions he is attempting to compare. As stated before, we are uninterested in what aspects of Joe's reasoning may or may not have been canonically correct. Rather, we chose to look at his reasoning in detail as an example of a student who started with an apparently weak and vague notion of Taylor series who then went on to refine it into a more robust way of reasoning with the topic, as played out through revisions of his graphical image.

Additionally, what we hope to have showed here is one example of a way that looking at a student's *process* of considering and seeking additional evidence can highlight how his graphical image changed over time. That is, examining Joe's reasoning on a very fine-grained level can say more than simply what he understands about the particular topic of Taylor series. It provides a lens into the ways that he visually played out his changing notions of what he means by "confidence interval" in such a way that permitted him to more successfully engage with the content of Tasks 1 and 2. Through his reasoning, we see that the notion of confidence interval, despite its flaws, proved to be a robust construct by which he was able to amend his graphical image toward a more appropriate, normative, and flexible understanding than one would imagine, given the initial state of his reasoning about Task 1. Additionally, through the evolution of his graphical image, we have a referent for the vague and/or confusing language that may be associated with his notion of confidence interval.

Contributions

As Borgen and Manu (2002) emphasize, "an understanding of what images, both correct and incorrect, that students might construct is important if teachers are to help students work toward connected formalizations" (p. 164). Even better – knowing *how* students build those images provides additional perspectives for informing pedagogy around the topic of Taylor series. Returning to Martin's (2009) point, recognizing that graphical representations of Taylor series are one of the most significant factors in separating novices from experts, it is instructive to work on building students' graphical images for such a topic. However, one cannot responsibly undertake that task without first exploring how students create that understanding for themselves, studying how they leverage their prior knowledge and reasoning on a moment-to-moment basis, building their concept of Taylor series from the many related, and more elementary, concepts that come before it.

References

Alcock, L. & Simpson, A. (2004). Convergence of Sequences and Series: Interactions Between Visual Reasoning and the Learner's Beliefs about their own Role. *Educational Studies in Mathematics, 57,* 1-32.

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.

Borgen, K. & Manu, S. (2002). What do students really understand? *Journal of Mathematical Behavior, 21*, 151-165.

Calais, G. (2008). Microgenetic analysis of learning: Measuring change as it occurs. *National Forum of Applied Educational Research Journal, 21*(3).

Habre, S. (2009). Multiple representations and the understanding of Taylor polynomials. *PRIMUS 19*(5), 417-432.

Kindfield, A. (1993). Biology diagrams: Tools to think with. *The Journal of the Learning Sciences, 3*(1), 1-36.

Martin, J. (2009). Expert conceptualizations of the convergence of Taylor series: Yesterday, today, and tomorrow. Doctoral Dissertation, University of Oklahoma.

Martin, J., Oehrtman, M., Roh, K. H., Swinyard, C. & Hart-Weber, C. (2011). Students' reinvention of formal definitions of series and pointwise convergence. In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education, Vol. 1, pg 239-254. Portland, Oregon.

Monaghan, J. (1991). Problems with the language of limits. *For the Learning of Mathematics, 11*(3), 20-24.

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.

Parnafes, O. (2009). Self-generated representations (SGR) for promoting conceptual understanding. *Proceedings of the Chais conference on instructional technologies research 2009: Learning in the technological era*. Y. Eshet-Alkalai, A. Caspi, S. Eden, N. Geri, Y. Yair (Eds.), Raanana: The Open University of Israel.

Schoenfeld, A., Smith, J. & Arcavi, A. (1993). Learning. In R. Glaser (Ed.), *Advances in Instructional Psychology* (Vol. 4, pp. 55-175). Mahwah, NJ: Erlbaum.

Sherin, B., Krakowski, M, & Lee, V. (2012). Some assembly required: How scientific explanations are constructed during clinical interviews. *Journal of Research in Science Teaching*, *49,* 166-198.

Soto-Johnson, H. (1998). The Impact of Technology on Infinite Series, *International Journal of Computer Algebra in Mathematics Education, 5*(2), p. 95-109.

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.

Yerushalmy, M., Schwartz, J.L. (1999) A procedural approach to exploration in calculus. *International Journal of Mathematical Education in Science and Technology*, 30 (6), 903-914.

THE STATUS OF CAPSTONE COURSES IN THE PREPARATION OF SECONDARY MATHEMATICS TEACHERS

Joshua Chesler	Dana Cox	Mary Beisiegel
CSU, Long Beach	Miami University	Harvard GSE
Rachael Kenney	Jill Newton	Jami Stone
Purdue University	Purdue University	Black Hills State University

For more than a decade, capstone courses have been recommended as a way for pre-service secondary mathematics teachers to connect the mathematics they learn in college to the mathematics they will teach in their own classrooms. Yet little is known about the extent and nature of the implementation of these courses in the United States. This paper presents findings from a 2011 survey of U.S. colleges and universities that investigated whether and how capstone courses for pre-service secondary mathematics teachers have been implemented.

Keywords: capstone course, mathematics teacher preparation, secondary mathematics

In 2001, the Conference Board of the Mathematical Sciences (CBMS) recommended that pre-service high school teachers complete "a 6-hour capstone course connecting their college mathematics courses with high school mathematics" (p. 8). Since that time, there have been a handful of reports on implementations of individual courses that fit this description (e.g., Artzt, Sultan, Curcio, & Gurl, 2011; Hill & Senk, 2004; Loe & Rezak, 2006; Shoaf, 2000; Van Voorst, 2004). However, the status of the mathematics capstone course in the United States is largely unknown. There has, thus far, been no systematic study of the extent or characteristics of its varied implementations.

Herein, we present results from a 2011 survey of colleges and universities that may offer an upper-level capstone course, either in the mathematics department or in the college of education, for mathematics majors intending to be secondary teachers. The goal of the survey was to investigate the status of capstone courses in the United States and the extent to which the CBMS recommendations align with the capstone courses in our sample. For the purposes of the survey, we defined a *capstone* as a course taken at the conclusion of a program of study for pre-service secondary mathematics teachers that places a primary focus on providing at least one of the following: (1) bridges between upper-level mathematics courses, (2) connections to high school mathematics, (3) additional exposure to mathematics content in which students may be deficient, or (4) experiences communicating with and about mathematics (Loe & Rezac, 2006).

The survey, which can be found in Appendix A, investigated the prevalence and nature of courses fitting this description. In particular, the survey included questions about capstone characteristics such as the department, title, duration, textbook(s), and other resources used in the course. It also included questions related to the nature of the course; specifically, data was collected about the description of the capstone course in the university's catalog, the course goals, the instructional style, and the content. To provide a more complete picture of the current state of capstone courses, data was also collected about instructors' backgrounds and their levels of academic freedom.

Perspective

Secondary mathematics teacher preparation programs typically require pre-service teachers to complete a mathematics major, or the equivalent (Artzt, Sultan, Curcio, & Gurl, 2011; CBMS, 2001). However, there is some uncertainty about the value of a traditional mathematics undergraduate degree for secondary mathematics teachers. The CBMS (2012), echoing the concerns of Felix Klein (1932), described a “double discontinuity” often encountered by secondary mathematics teachers. The first is when they transition from high school mathematics to seemingly disconnected university mathematics courses. The second occurs when new teachers, upon beginning their careers, experience a disconnect between the mathematics learned in university courses and the mathematics of high school. These ideas align with Monk’s (1994) influential report which placed doubt on the value of the upper division mathematics courses for preparing effective mathematics teachers. Among the conclusions from the large-scale longitudinal study, Monk declared that “having a mathematics major has no apparent effect on student performance” (p. 132).

Hodgson (2001) noted that pre-service secondary mathematics teachers “have no explicit occasion for making connections with the mathematical topics for which they will be responsible in school, nor of looking at those topics from an advanced point of view” (p. 509). He endorsed the inclusion of undergraduate coursework to help pre-service teachers develop “deep conceptual understanding of the school mathematics content” (p. 512). The CBMS (2001) recommendation for capstone courses arose from a similar recognition that an undergraduate degree in mathematics may not help pre-service teachers develop this deep and relevant knowledge prior to entering their profession. A decade after this recommendation, the survey reported herein provides insight about the status of the capstone course for pre-service secondary mathematics teachers.

Methodology

From the 1,713 institutions listed by the Carnegie Foundation for the Advancement of Teaching (Carnegie Classifications, 2011), we selected a stratified random sample of 200 institutions, weighted appropriately for each of nine classification groups (e.g., PhD granting institutions with high research activity, Master’s Colleges and Universities-larger programs). A 23 question survey (see Appendix A) was developed using Qualtrics online survey software and sent to each of these 200 institutions. The first two questions (P1 & P2) inquired about whether the institution has a capstone course. Institutions with capstones were then prompted to answer 21 additional questions (Q1 though Q21). As only 32 of these 200 institutions responded, the sample was expanded to a total of 73 by sending the survey to three relevant email listservs. This second phase of solicitation altered our initial plan for random sampling; our ability to make inferences has, thus, been hindered. However, the sample provided rich data which was analyzed in Excel using basic summative statistics. The responses for each of the 21 survey questions were analyzed separately by at least two team members. The analyses were then compared, merged, and summarized by the research team.

Results

The survey was completed by individuals at 73 distinct colleges and universities. Of these institutions, 42 (57.5%) reported having a content course, taken at the conclusion of a program of study for pre-service secondary mathematics teachers, that satisfies at least one of the goals that Loe & Rezac (2006) described for their capstone course. That is, each of the 42 institutions has a course intended to provide at least one of the following:

1. bridges between upper-level mathematics courses,
2. connections to high school mathematics,
3. additional exposure to mathematics content in which students may be deficient, or
4. experiences communicating with and about mathematics.

The respondents represented a variety of institutions, as reflected in the 2011 Carnegie classifications; this data is summarized, along with additional information, in Table 1.

Among the 42 institutions reporting capstone courses, one submitted separately about two different courses, and two did not provide any additional details about their courses.

Table 1. Summary of the sample.

Carnegie Type	All	Have Cap-stone	CBMS
Bac/A&S: Baccalaureate Colleges--Arts & Sciences	12	6	0
Bac/Assoc: Baccalaureate/Associate's Colleges	3	1	0
Bac/Diverse: Baccalaureate Colleges--Diverse Fields	8	7	6
Master's L: Master's Colleges and Universities (larger programs)*	24	15	10
Master's M: Master's Colleges and Universities (medium programs)*	7	5	3
Master's S: Master's Colleges and Universities (smaller programs)	4	3	2
DRU: Doctoral/Research Universities	3	0	0
RU/H: Research Universities (high research activity)	3	3	3
RU/VH: Research Universities (very high research activity)	8	2	2
Spec/Faith: Special Focus Institutions--Theological seminaries, Bible colleges, and other faith-related institutions	1	0	0
TOTAL	73	42	26

* Each of these categories has one respondent that has a capstone but did not answer follow-up questions; it is unknown whether they align with the CBMS recommendation.

CMBS versus non-CBMS. As our survey defined a capstone course more broadly than the CBMS recommendation, most of the results reported below make a distinction between what we have labeled as CBMS and non-CBMS courses. A CBMS course is one that aligns with the CBMS recommendation of "connecting [students'] college mathematics courses with high school mathematics" (2001, p.8). By parsing the data in this way, we were able to separately comment on the statuses of capstone courses which align with the CBMS recommendation and those self-identified capstones which do not. This criteria was operationalized in question Q7 (see Appendix A), which investigated the purposes of the course. Table 2 summarizes responses to Q7 and lists the six capstone course purposes which followed Q7's instructions to "check all that apply." The first four purposes align with capstone goals enumerated by Loe and Rezac (2006). Capstone courses which had purpose (b) were classified here as CBMS courses; a non-CBMS course is one which aligns with at least one of the other purposes, but does not align with purpose (b). There were 41 responses about capstone course goals; of these 41 capstones, 26 were categorized as CBMS courses. A mean of 3.2 goals were chosen per course.

Table 2. Purpose of the capstone (n=41).

Purpose of capstone is to provide:	n	%
(a) bridges between upper-level mathematics courses, especially real analysis, abstract algebra, probability/statistics, and geometry	22	54%
(b) an opportunity to explore connections between college mathematics and secondary school mathematics	26	63%
(c) additional exposure to areas of mathematics in which they may be deficient	24	59%
(d) research and writing in mathematics and with making oral presentations to their peers and instructors	33	80%
(e) the opportunity to learn pedagogical principles for teaching secondary mathematics	9	22%
(f) opportunities to become familiar with technology for teaching	9	22%
(g) other	8	20%

Capstone Course Goals. The CBMS vs. non-CBMS distinction was apparent in the results of question Q8 which investigated the goals of the capstone courses (see Table 3). Goals (b) and (e) in Table 3 correspond to the CBMS recommendations and were much more prevalent in the CBMS courses. The most common goal for both CBMS and non-CBMS courses was for students to develop a deeper understanding of mathematics. Survey respondents were given an opportunity to name goals that were not given in the survey list. Examples of non-CBMS goals included student investigation of a substantial mathematics topic and learning advanced mathematics on their own, while an example of a CMBS goal was clearly writing mathematics.

Table 3. Goals of capstones.

Goals	All	CBMS	non-CBMS
(a) Students are knowledgeable about the university mathematics content addressed in the course	56%	50%	67%
(b) Students take an in-depth look at some mathematical topics which are particularly important in secondary mathematics	56%	77%	20%
(c) Students know how to use a variety of teaching strategies when teaching mathematics	15%	23%	0%
(d) Students can (effectively) integrate technology into their future classrooms	24%	35%	7%
(e) Students connect appropriate college mathematics content to high school mathematics content and pedagogy	46%	69%	7%
(f) Students become aware of current topics and issues in secondary school mathematics	17%	23%	7%
(g) Students develop a deeper appreciation of mathematics	85%	81%	93%
(h) Students develop a personal philosophy to support the teaching of secondary mathematics	20%	27%	7%
(i) Other	20%	8%	40%

n	41	26	15

Length of the capstone course. The majority of capstone courses in our sample were offered as one-semester or one-quarter classes. A larger proportion of CBMS capstone courses (73%) were single courses, whereas half of non-CBMS capstone courses spanned more than one semester. Data about the number of capstone semesters/quarters are summarized in Table 4. The survey also revealed a wide range of times since the capstone was first offered at the institutions in the sample, from one to more than twenty years. Across all institutions, the capstone courses had existed for a median of seven years. For CBMS capstone courses, the median length of existence was six years; non-CBMS capstone courses existed for a median of ten years.

Table 4. Number of capstone semesters/quarters (n=40 institutions).

# of courses	All	CBMS	non-CBMS
1	26	19	7
2	11	5	6
3	2	1	1
4	1	1	0
Total	40	26	14

Capstone course resources. The resources used to develop the courses are summarized in Table 5. On average, CBMS capstone courses were developed in consultation with three of the listed resources, where non-CBMS capstone courses were developed with a mean of 1.5 resources. The development of CBMS courses was, to a much larger extent, guided by national organizations and recommendations, as well as by high school standards. Four courses (three CBMS) were developed in consultation with education departments; other departments consulted were communications (CBMS) and science departments (non-CBMS).

Table 5. Resources used to develop course (n=41 capstones).

Resources used to develop course	All	CBMS	non-CBMS
National guidelines	13	11	2
Common Core State Standards	9	9	0
National Council of Teachers of Mathematics	17	17	0
Conference Board of the Mathematical Sciences	11	8	3
Mathematics Association of America	22	14	8
National Mathematics Advisory Board Recommendations	6	4	2
Collaboration with other departments on campus	8	6	2
Collaboration with other universities	5	4	1

Capstone course students. Twelve capstone courses, all of which were CBMS courses, were described as being required specifically for pre-service mathematics teachers. At the non-CBMS schools, all of the students who enrolled in the courses were mathematics majors. At most schools (both CBMS and non-CBMS), students intending to be mathematics teachers did not exclusively populate the capstone courses. Indeed, only six capstone courses (all

CBMS) reported that they are exclusively for students seeking teaching licensure. Two of the non-CBMS courses did not include any category of students seeking licensure. Table 6 lists the percentages of capstone courses in our sample that included various categories of students.

Table 6. Students to whom the capstone is available.

Who takes the course?	**All**	**CBMS**	**non-CBMS**
Alternate licensure students post-baccalaureate	22%	31%	7%
Graduate students	10%	12%	7%
Undergraduate math majors	80%	69%	100%
Undergraduate math majors pursuing teaching licensure	83%	85%	80%
Undergraduate mathematics education majors pursuing teaching licensure	59%	65%	47%
Undergraduate math minors	34%	31%	40%
Undergraduate math minors pursuing licensure	27%	31%	20%
n	41	26	15

Capstone course prerequisites. Our expectation was that the capstone course, as defined in this survey, is typically intended to be taken at the conclusion of a program of study for pre-service secondary teachers. Therefore, our survey probed the prerequisites for these courses. Five responses stated only that advanced standing was required; these responses have been eliminated from

Table 7, which provides details about prerequisites. Calculus and linear algebra were the most commonly listed prerequisites. The one capstone course which did not include calculus as a prerequisite required a mathematics course specifically for pre-service mathematics teachers and six additional units of unspecified mathematics.

Some features of this list of prerequisites stand out, particularly when comparing CBMS to non-CBMS courses. The CBMS capstone courses were twice as likely to have non-Euclidean (rather than Euclidean) geometry as a prerequisite. These two geometry courses were equally likely prerequisites among the non-CBMS courses. Calculus-based statistics was more popular as a prerequisite among CBMS courses; eight of the nine non-CBMS courses which required statistics did not require it to be calculus-based. If Probability, Calculus-Based Statistics, Non-Euclidean Geometry, Abstract Algebra, and Real Analysis are counted as upper-division courses, then 31% of all capstone courses reported no upper division prerequisites. This rate was consistent among both CBMS and non-CBMS courses, though there is divergence when higher numbers of upper division prerequisites are considered. Among the CBMS courses, 65% required two or fewer upper division prerequisites, while only 46% of non-CBMS courses required two or fewer.

Table 7. Prerequisites for the capstone.

Course Name	All	All %	CBMS	CBMS %	Non-CBMS	non-CBMS %
Calculus	35	97%	22	96%	13	100%
Linear Algebra	31	86%	18	78%	13	100%
Discrete Mathematics	6	17%	4	17%	2	15%
*Abstract Algebra	14	39%	9	39%	5	38%
Euclidean Geometry	13	36%	6	26%	7	54%
*Probability	9	25%	7	30%	2	15%
*Real Analysis	15	42%	12	52%	3	23%
*Calculus-Based Statistics	8	22%	7	30%	1	8%
Other	18	50%	11	48%	7	54%
Statistics with no Calculus prereq.	15	42%	7	30%	8	62%
*Non-Euclidean Geometry	19	53%	12	52%	7	54%
Combinatorics	14	39%	9	39%	5	38%
n	36		23		13	

* Upper-division courses

Capstone course instruction and content. Survey respondents were asked to describe the academic background of the instructor who has most often taught the course in the past five years. Table 8 summarizes the results. At least 14 out of 15 non-CBMS course instructors had backgrounds in mathematics; the fifteenth capstone course was reported to be conducted with individual instructors paired with students. One CBMS capstone course was co-taught by a mathematician and mathematics educator. Only four instructors, all of whom teach CBMS capstones, were reported to exclusively have a mathematics education background.

Table 9. Instructor backgrounds.

Instructor Background	All	CBMS	non-CBMS
Mathematics	35	21	14
Mathematics Education	14	12	2
Both Math & Math Ed	10	8	2
n	41	26	15

Note: Some capstones instructor backgrounds are not reflected in this table. There was one CBMS course instructor with a computer science background. One non-CBMS capstone course paired individual students and faculty members. Some instructors selected multiple backgrounds.

Survey respondents were also asked to comment on the level of instructor freedom in choosing the topics examined in the capstone course. Thirty-three of 41 capstone courses (80.5%) selected the following: “A lot - There are limited guidelines or recommendations for teaching this course, so instructors get to choose the materials they want to use.” The rate was consistent across CBMS and non-CBMS courses. Only one capstone course (CBMS) was reported to have no instructor freedom because a course coordinator chooses the materials. The other seven courses had some instructor freedom in the choice of topics; their chosen survey option was, “there are recommended curriculum materials, but the instructor is not required to use them.” The survey also investigated instructor freedom in how the course was taught or structured. For this question, 100% of respondents reported yes to one of the following choices:

- Some - the department has recommendations for how the class is taught and expects instructors to use those recommendations as a guide, but not an imperative. (n=15)
- A lot - the department has no recommendations for how the course should be taught, so it is up to the instructor to decide how to teach the course. (n=25)

This level of instructor freedom was reflected in the variety of materials used for the courses. Among the 31 responses to questions about course materials, 18 different books were listed as course textbooks, 13 courses used various materials, and at least four used materials primarily developed internally. Among the many texts listed, only three were listed as a textbook for three or more capstone courses:

- *Mathematical connections: A capstone course*. Conway, J. (2010) – 3 courses
- *Mathematics for high school teachers: An advanced perspective.* Usiskin, Peressini, & Marchisotto, & Stanley (2002) – 7 courses
- *The mathematics that every secondary school math teacher needs to know.* Sultan & Artzt (2010) – 3 courses

Likewise, a wide variety of classroom technologies were used in the capstone courses. Of 39 respondents on this topic, only two reported to not use any technology in the course (both were non-CBMS courses). The most commonly used tools were Geometer’s Sketchpad or Geogebra (15 and 6, respectively), graphing calculators (21), and Microsoft Excel (16). There was not a pronounced difference between CBMS and non-CBMS courses other than in the use of Excel; all 16 of the capstones that used Excel were CBMS courses.

Variety was detected in the content of the capstone courses. A survey question asked, “In the last semester that the course was taught, what mathematical or pedagogical topics were examined?” Table 10 shows counts for some categories of responses to this question. As compared with the non-CBMS courses, the CBMS courses included more secondary mathematics topics and pedagogical concerns. All of the non-CBMS courses addressed advanced mathematical topics.

Table 10. Categories of topics covered.

Topic	All	CBMS	Non-CBMS
Deeper look at secondary mathematics	11	11	0
Advanced mathematical topics	22	10	12
History of mathematics	7	4	3
Pedagogical concerns	6	6	0
n	33	21	12

In a typical semester or quarter, more than 60% of class time was spent on a combination of whole-class discussion, students working with partners or in small groups, and students working independently. The percentage of time devoted to each of these types of student work varied between CBMS and non-CBMS capstone courses. Notably, non-CBMS capstone courses devoted a larger amount of class time to students working independently (41% vs. 23% for CBMS capstones). Among all capstones, lectures accounted for 18% of class time

(16% for CBMS, 21% for non-CBMS capstones). The percentages of class time associated with different lesson implementations are summarized in Table 10.

Table 10. In a semester/quarter, percentages of class time spent using various lesson designs/implementations.

Percentage of class time spent on:	All	CBMS	non-CBMS
Lecture	18%	16%	21%
Whole-class discussion	20%	26%	12%
Students working with partners or in small groups	17%	20%	10%
Students working independently	30%	23%	41%
Students exploring mathematical concepts using manipulatives	3%	4%	1%
Students exploring mathematical concepts using technology	4%	5%	1%
Student Presentations	7%	5%	11%
Other	1%	0%	3%
n	41	26	15

Among the capstone courses surveyed, tests, presentations, and the reading of articles were reported as the most popular type of assignments. Each of these assignments, however, was more popular in CBMS capstone courses than in non-CBMS. Table 11 lists the percentages of respondents who use each of the listed assignments or activities.

Table 11. Major assignments and in-class activities.

Assignments/Activities	All	CBMS	non-CBMS
Portfolios of course reflections	20%	23%	13%
Plan and present lessons to the class	39%	54%	13%
Plan and present lessons to secondary school mathematics classes	10%	15%	0%
Analyze K - 12 textbooks and curriculum materials	12%	19%	0%
Read and report on articles from practitioner journals	34%	38%	27%
Field placements	2%	4%	0%
Classroom Observations	10%	15%	0%
Tests/quizzes	32%	38%	20%
n	41	26	15

Discussion

In 2001, the CBMS recommended that pre-service high school mathematics teachers complete "a 6-hour capstone course connecting their college mathematics courses with high school mathematics" (p. 8). Ten years later, courses which align with this recommendation seem not to be wide-spread. Only 26 out of the 73 institutions in our survey had at least one course which aligns with the CBMS recommendation. Furthermore, assuming that six hours of coursework would span more than one semester/quarter, only 7 of the 26 CBMS capstone courses in our sample likely satisfy this requirement. Looking beyond the CBMS recommendation, 16 additional institutions in our sample provide a capstone experience (not aligned with the CBMS) for this population of students.

The CBMS vs. non-CBMS distinction was determined by the stated purposes of the capstone course. A CBMS capstone course has the (not necessarily sole) purpose to connect college and high school mathematics, as recommended by the CBMS. Our survey, however, used a broader definition of capstone and included courses which fostered connections between college-level courses, provided exposure to additional mathematics content, and/or engaged students in communicating with or about mathematics. Indeed, most capstone courses reported in our survey addressed many of these and other goals and served multiple purposes. Our survey data indicates diversity across many characteristics of the courses which respondents identified as capstones.

Despite this diversity, some general features are shared by most capstone courses in our sample. These courses integrate group or individual student coursework during class time; on average, only 18% of time is devoted to lecture. The use of (not necessarily instructional) technology was popular among nearly all of the courses. All 41 capstone courses were completed by pre-service secondary mathematics teachers at the end of their undergraduate experience; however, only 12 of the 41 capstone courses were taken exclusively by pre-service secondary teachers. This lack of exclusivity may be connected to the CBMS observation that courses for future teachers may be difficult to implement in institutions that serve a small number of pre-service mathematics teachers (CBMS, 2012). Our survey, however, did not reveal this level of detail. In general, instructors reported a large amount of freedom in choosing the content and instructional style for their courses. This freedom is also reflected in the wide variety of assessment devices and resources used. It is possible that this is a byproduct of the capstone being a relatively new type of course. Indeed, a defining feature of the current state of capstone courses is the variety of implementations.

Within this variety, there are notable differences between CBMS and non-CBMS courses. As would be expected given the recentness of the CBMS recommendation, non-CBMS courses are typically older than the CBMS courses (10 vs. 7 years in median time since first offered). Furthermore, CBMS capstones are more likely to have been developed in consultation with national guidelines from mathematics and educational organizations. They are also more likely to be taught by someone with a mathematics education background. Though most (69%) capstone courses required upper division courses as pre-requisites, there were some differences in the type of courses required by CBMS courses, particularly in the areas of geometry and statistics and in the quantity of upper division prerequisites (more were required by non-CBMS capstones).

Given these differences between the two categories of capstones, along with their differences in purpose, it would be tempting to characterize the differences between CBMS and non-CBMS courses as being signs of different programmatic foci. Specifically, perhaps the CBMS courses are located in programs more focused on teacher preparation. However, there are also signs which indicate that this may not be the case. Notably, CBMS courses are more likely to include a calculus-based statistics course (instead of a lower-level statistics course) as a prerequisite and are less likely to have a Euclidean geometry prerequisite. That is, the prerequisite coursework in programs with CBMS capstones may be less amenable to making connections to high school content throughout the undergraduate program. Indeed, a capstone which focuses on high school connections may be more of a necessity in departments with prerequisite coursework which does not support this. The nature of an individual capstone course may indicate little about the program which houses it.

In February 2012, the CBMS released a draft of an update to their 2001 recommendations for the mathematics education of teachers (CBMS, 2012). The new document does not include the word "capstone." Instead, the CBMS recommends that pre-service secondary mathematics teachers complete the equivalent of a mathematics major "that includes three courses with a primary focus on high school mathematics from an advanced viewpoint" (p. 7). Absent from the recommendations is advice on when these courses should be taken; in particular, there is no recommendation that these courses are intended as a capstone at the end of an undergraduate program. It has been barely more than a decade since the CBMS recommended the capstone and, though the recommendation was not renewed in the 2012 draft, the CBMS has strengthened the recommendation for pre-service teachers to interact with high school mathematics content at a deeper level. Though our study was more widely focused than trying to measure the impact of the CBMS recommendation, the survey results give some indication of how the new recommendations may be interpreted and implemented. More generally, though, our survey uncovered and described much about the status of capstone courses in the preparation of secondary mathematics teachers.

References

Artzt, A., Sultan, A., Curcio, F. R., & Gurl, T. (2011). A capstone mathematics course for prospective secondary mathematics teachers. *Journal of Mathematics Teacher Education*, Online First, June 14, 2011.

Carnegie Classifications. (2011). *Carnegie Foundation for the Advancement of Teaching*. Retrieved from http://classifications.carnegiefoundation.org/resources/

Conference Board of the Mathematical Sciences (CBMS). (2001). *The mathematical education of teachers.* Providence, RI: American Mathematical Society.

Conference Board of the Mathematical Sciences (CBMS). (2012). *The mathematical education of teachers II: Draft for public discussion, February 9, 2012.* Retrieved from http://www.cbmsweb.org/MET2/MET2Draft.pdf

Conway, J. B. (2010). *Mathematical connections: A capstone course*. Providence, RI: American Mathematical Society.

Hill, R., & Senk, S. (2004). A capstone course for prospective high school mathematics teachers. *Mathematicians and Education Reform Newsletter*, *16*(2), 8-11.

Hodgson, B. (2001). The mathematical education of school teachers: Role and responsibilities of university mathematicians. In D. A. Holton (Ed.) *The teaching and learning of mathematics at the university level: An ICMI study* (pp. 501-518). Boston, MA: Kluwer Academic Publishers.

Klein, F. (1932). *Elementary mathematics from an advanced standpoint: Arithmetic, algebra, analysis.* New York: Macmillan.

Loe, M., & Rezak, H. (2006). Creating and implementing a capstone course for future secondary mathematics teachers. In K. Lynch-Davis and R. L. Rider (Eds.) *The work of mathematics teacher dducators: Continuing the conversation Monograph Series Volume 3*(pp. 45-62). San Diego, CA: AMTE.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of education review*, *13*(2), 125-145.

Shoaf, M. M. (2000). A capstone course for pre-service secondary mathematics teachers. *International Journal of Mathematical Education in Science and Technology*, *31*(1), 151-160.

Usiskin, Z., Peressini, A. L., Marchisotto, E., & Stanley, D. (2002). *Mathematics for high school teachers: An advanced perspective* (1st ed.). Upper Saddle River, NJ: Prentice Hall.

Van Voorst, C. (2004). Capstone mathematics course for teachers. *Issues in the Undergraduate Mathematics Preparation of School Teachers: The Journal*, *4*, 1-11.

Sultan, A., & Artzt, A. F. (2011). *The mathematics that every secondary school math teacher needs to know*. New York, NY: Routledge.

Appendix A: Survey Questions

P1. What is the name of your institution? (This will only be used internally and will be removed from the data during analysis.)

P2. Capstone Course Definition: For the purpose of this survey, we define a capstone course for pre-service secondary mathematics teachers in the following way: A Capstone course is a content course taken at the conclusion of a program of study for pre-service secondary mathematics teachers that satisfies at least one of the following criteria for a capstone course (Loe and Rezac, 2006): (1) Provide bridges among upper-level mathematics courses, especially real analysis, abstract algebra, and geometry; (2) Provide preservice teachers an opportunity to explore connections to the high school curriculum so that they have a better understanding of the mathematics they will teach; (3) Provide preservice teachers with additional exposure to areas of mathematics in which they may be deficient; (4) Provide preservice teachers experiences with research and writing in mathematics and oral presentations to their peers and instructors. Please exclude from this definition a course that is specifically related to mathematics teaching methods (i.e., a "methods" course). Based on this definition, does your department offer at least one capstone course to pre-service secondary mathematics teachers?

Q1. What is the name of the course(s)?

Q2. How many total credit hours are offered for the course(s)?

Q3. How are the total credit hours divided among different forms of the class, such as some hours of lecture and some hours of lab/workshop or practicum? Please fill in the number of hours for each below:
(a) Lecture, (b) Workshop/Lab/Activity Hours, (c) Practicum Hours, (d) Other

Q4. How long is the duration of the course (i.e., the number of quarters, semesters, or years)?

Q5. How is this course described to students? (If your institution has an on-line course catalog, it would be acceptable to copy and paste the description here.)

Q6. How long has this / course been offered at your institution? Time (in years):

Q7. What purpose does the capstone course serve in your program of study? Please check all that apply:
(a) To provide bridges between upper-level mathematics courses, especially real analysis, abstract algebra, probability/statistics, and geometry, (b) To provide pre-service teachers with an opportunity to explore connections between college mathematics and secondary school mathematics, (c) To provide pre-service teachers with additional exposure to areas of mathematics in which they may be deficient, (d) To provide pre-service teachers experiences with research and writing in mathematics and with making oral presentations to their peers and instructors, (e) To provide pre-service teachers with the opportunity to learn pedagogical principles for teaching secondary mathematics, (f) To provide pre-service teachers with opportunities to become familiar with technology for teaching, (g) Other (please describe)

Q8. What outcomes/goals do you have for students enrolled in your capstone course? Please check all that apply:

(a) Students are knowledgeable about the university mathematics content addressed in the course, (b) Students take an in-depth look at some mathematical topics which are particularly important in secondary mathematics, (c) Students know how to use a variety of teaching strategies when teaching mathematics, (d) Students can (effectively) integrate technology into their future classrooms, (e) Students connect appropriate college mathematics content to high school mathematics content and pedagogy, (f) Students become aware of current topics and issues in secondary school mathematics, (g) Students develop a deeper appreciation of mathematics, (h) Students develop a personal philosophy to support the teaching of secondary mathematics, (i) Other

Q9. In a typical quarter or semester, what percentage of class time is spent engaging in the following activities?
(a) Lecture, (b) Whole-class discussion, (c) Students working with partners or in small groups, (d) Students working independently, (e) Students exploring mathematical concepts using manipulatives, (f) Students exploring mathematical concepts using technology, (g) Other, (h) Other-TEXT

Q10. What are some major assignments or in-class activities that are required of students in the capstone course offered at your university? Please check all that apply:
(a) Portfolios of course reflections, (b) Plan and present lessons to the class, (c) Plan and present lessons to secondary school mathematics classes, (d) Analyze K - 12 textbooks and curriculum materials, (e) Read and report on articles from practitioner journals, (f) Field placements, (g) Classroom Observations, (h) Tests/quizzes, (i) Other

Q11. What are the mathematics prerequisites for the capstone course at your university? Please check all that apply:
(a) Calculus, (b) Linear Algebra, (c) Combinatorics, (d) Probability, (e) Calculus-Based Statistics, (f) Statistics with no prerequisite in Calculus, (g) Euclidean Geometry, (h) Non-Euclidean Geometry, (i) Abstract Algebra, (j) Real Analysis, (k) Discrete Mathematics, (l) Other, (m) There are no prerequisites for the course

Q12A. Is this a required course for certain majors or degree options?

Q12B. To whom is this course available? Please check all that apply:
(a) Alternate licensure students post-baccalaureate, (b) Graduate students, (c) Undergraduate math majors, (d) Undergraduate math majors pursuing teaching licensure, (e) Undergraduate mathematics education majors pursuing teaching licensure, (f) Undergraduate math minors, (g) Undergraduate math minors pursuing licensure, (h) Other

Q13. In the past five years, how would you describe the academic background of the instructor who has most often taught your capstone course? Please check all that apply:
(a) Mathematics, (b) Mathematics Education, (c) Education, (d) Education Administration, (e) Curriculum and Instruction, (f) Other (please describe)

Q14. What titles of textbooks have been used by faculty or students in teaching the capstone course in the last five years? Please check all that apply:
(a) Bremigan, E., Bremigan, R. and Lorch, J. (2011). Mathematics for secondary school teachers. MAA., (b) Cooney, T.J., Brown, S.I., Dossey, J.A., & Wittmann, E.Ch. (1996). Mathematics, pedagogy, and secondary teacher education. Portsmouth, NH: Heinemann., (c) Conway, J. (2010). Mathematical connections: A capstone course. Providence, RI: AMS., (d) Cuoco, A. (2005). Mathematical connections: A companion for teachers and others. Newton, MA: Educational Development Center., (e) Sultan, A., & Artzt, A.F. (2010). The mathematics that every secondary school math teacher needs to know. Hoboken, NJ: Taylor & Francis., (f) Usiskin, Z., Stanley, R., Peressin, A., & Marchisotto, E. (2003). Mathematics for high school teachers: An advanced perspective. Needham, MA: Prentice Hall., (g) Other (ISBN would suffice)

Q15. Has your department or have instructors of this course developed any supplemental materials for this course? Describe any additional materials have you developed or incorporated in this course

Q16. In the last semester that the course was taught, what mathematical or pedagogical topics were examined? If the capstone course changes from semester to semester, please indicate this along with a range of topics that you feel are representative of those included.

Q17. How much instructor freedom is permitted in choosing the topics examined in the capstone course?
(a) None - the course has a coordinator that chooses all of the materials/textbook, etc., (b) None - the course was developed by a curriculum committee., (c) Some - there are recommended curriculum materials, but the instructor is not required to use them., (d) A lot - There are limited guidelines or recommendations for teaching this course, so instructors get to choose the materials they want to use., (e) Please comment as needed

Q18. How much instructor freedom is permitted for how the class is taught/structured?
(a) None - the department has recommendations for how the class is taught and expects instructors to closely follow those recommendations., (b) Some - the department has recommendations for how the class is taught and expects instructors to use those recommendations as a guide, but not an imperative., (c) A lot - the department has no recommendations for how the course should be taught, so it is up to the instructor to decide how to teach the course., (d) Please comment as needed

Q19. What resources were used to develop this course? Please check all that apply:
(a) State guidelines, (b) National guidelines, (c) Common Core State Standards, (d) National Council of Teachers of Mathematics, (e) Conference Board of the Mathematical Sciences, (f) Mathematics Association of America, (g) National Mathematics Advisory Board Recommendations, (h) Collaboration with other departments on campus (please name those departments), (j) Collaboration with other universities, (k) Other (please describe)

Q20. Check any classroom technology used in this course that these future teachers may eventually use in their classrooms?
(a) Geometer's Sketchpad, (b) Geogebra, (c) Graphing Calculators, (d) Excel, (e) Fathom, (f) TinkerPlots, (g) Handheld devices, (h) Cell phones/applications, (i) Wikis/ Social Networking Tools, (j) Other, (k) No technology is used

Q21. Have you done any follow-up on the usefulness/success of this course? You will not be asked to describe the outcomes. Check all that apply:
(a) End of course teacher evaluation, (b) Exit interviews with students, (c) Feedback from student teachers, (d) Longitudinal research on this course, (e) Anecdotal evidence on the effectiveness of the course, (f) Other:, (h) No follow-up data has been collected

Q22. If available, please take a moment to upload a recent syllabus used in your capstone course. Microsoft Word or PDF formats are all fine (.doc, .docx, .pdf).

TEACHING METHODS COMPARISON IN A LARGE INTRODUCTORY CALCULUS CLASS

Warren Code, David Kohler, Costanza Piccolo, Mark MacLean
University of British Columbia

We have implemented a classroom experiment similar to a recent study in Physics (Deslauriers, Schelew, & Wieman, 2011)*: each of two sections of the same Calculus 1 course at a research-focused university were subject to an "intervention" week where a less-experienced instructor encouraged a much higher level of student engagement by design; we employed a modified quasi-experiment structure for our methods comparison with a Calculus 1 student population and with further steps to improve validity. Our instructional choices encouraged active learning (answering "clicker" questions, small-group discussions, worksheets) during a significant amount of class time, building on assigned pre-class tasks. The lesson content and analysis of the assessments were informed by existing research on student learning of mathematics, in particular the APOS framework. We report improved student performance, on conceptual items in particular, in the higher engagement section in both cases.*

Key words: Calculus, design experiment, classroom experiment

Introduction and Research Questions

Motivation for this study has both local and general components. At the research-focused institution in question, the calculus course under study has, in terms of mathematical background and interest, a diverse audience consisting primarily of Commerce and Economics students. Though it is terminal for many of the commerce students, the course is not a fully separate "Business Calculus" course in that it is equivalent in course credit to the more traditional, science-oriented calculus course taught by the same department; as a result, much of the syllabus is in common with Calculus 1 elsewhere in North America. Traditional lecture remains the (nearly) uniform choice for instruction, though some local pressure to examine teaching methods has recently arisen:

- Both the School of Business and Economics Department are interested in improving students' skills at dealing conceptually with calculus in the context of other courses;
- Student demand for high levels of classroom engagement has been increasing at the institution; and
- The Mathematics Department has identified a wide spread of performance in this course and wishes to address student difficulties.

On a broader scale, our work is motivated by a demand for empirical study of less-traditional but evidence-based instructional methods for introductory calculus at the undergraduate level. We gleaned structural ideas from the Physics Education Research (PER) community, though instructional decisions in our study were based on research on students in mathematics, with an attempt to situate our analysis in the Action Process Object Schema (APOS) framework (Dubinsky & Mcdonald, 2002). Our research questions are not unlike those of Deslauriers et al. (2011):

Question 1: Compared to more traditional lecture-based instruction, will students demonstrate more sophisticated reasoning on an immediate test of learning when high-engagement instruction is implemented for a single topic (100-150 minutes of class time)?

Question 2: Will any effects persist to later, standard tests of learning in the course?

Theoretical Perspective

In seeking to introduce more extensive evidence-based teaching methods into the course, we considered existing models for instruction. Our lesson structures borrowed ideas from Peer Instruction (Crouch & Mazur, 2001) and general principles about learning that are now available (Bransford, Brown, & Cocking, 2003) but are not known to many university mathematics faculty, particularly at research-focused institutions. Specifically, the goal was to promote "active learning" as described in the science education literature, where much of the evidence arises from earlier stages of schooling but has seen some study at the post-secondary level (Hake, 1998; Michael, 2006). The key components of the instructional intervention were:

Pre-class activities:

- Students read and engage in structured exploration,
- Students submit responses online to specific questions about the material,
- Instructor browses responses prior to class time.

High-engagement class time:

- Group discussion and activities using structured notes (text with appropriate blank spaces to be filled in by student work) and worksheets,
- "Clicker" questions with follow-up discussion among students and/or whole-class directed by instructor,
- Reactive lecture based on comments or issues arising during class activities,
- Relatively small portion of time allocated to traditional, one-to-many exposition,
- Driven in part by pre-class results.

Identical online homework exercises were assigned to both sections after the instructional period, which was a similar arrangement to previous course years and the other topics. Student exposure in the interventions was thus largely compatible with the Activities, Class, Exercises (ACE) cycle (Weller et al., 2003).

Our study took place within the confines of the existing course material and expectations: students at this institution are expected to gain procedural fluency in calculus along with conceptual understanding, though the former tends to be assigned much more weight in assessment than the latter. We acknowledge that greater improvement may be possible by combining active learning pedagogy with a shift in course material to be more conceptual and more applied in nature and with more general program reform.

In designing lessons for the two classroom intervention topics, we considered sources in the literature for APOS-based study of both topics. For the first topic, related rates, we considered the work of Martin, (2000) and Engelke, (2007), which probed the combination of skills required for the solving of geometric related rates problems (i.e. involving a geometric relationship between quantities, like depth and volume of water in a draining tank of a certain shape), as well as the recent thesis of Tziritas, (2011), where a genetic decomposition for related rates problems was performed and tested.

For the second topic, linear approximation, we considered literature on covariational reasoning (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002) and students' relation of the tangent line to the graph of a function (Asiala, Cottrill, Dubinsky, & Schwingendorf, 1997).

Our quasi-experimental design extends that of Deslauriers et al. (2011), and is detailed in the Methodology section below. To our knowledge, and supported by a recent survey article (Speer, Smith III, & Horvath, 2010), no study of this kind has been reported for this size of college-level mathematics classroom.

Methodology

In this section, we describe the design of the quasi-experiment, including the setting, roles of the researchers involved, and the performance assessments implemented.

Outline of Experimental Design

We employed similar elements found in the Deslauriers et al. (2011) study:

- Natural setting of two similar large sections in the same course, during the same semester.
- Student performance measured early in the term to provide baseline data.
- Classroom intervention by an instructor with less experience but recent training on theories of learning and non-lecture pedagogy, replacing the assigned, well-regarded lecturer.
- Single topic intervention over approximately one week of classes.

We enhanced the experimental design in the following ways:

- Introducing a "crossover" by applying two single-topic interventions, one for each course section in a different week, to account for differing student populations.
- Removing the primary investigator further from the classroom intervention: though assisting in the development of instructional materials instruction, the primary investigator was not the instructor.
- Having the initial post-tests of learning based on agreed-upon learning objectives but written by someone not involved in the instructional design.
- Tracking student performance with respect to the two topics on subsequent course exams.

We expand on this outline in the remainder of this section. Figure 1 shows a timeline, including the positions of the common assessments.

Course Week	1	2…	7	8	9	10	11	12	end of term
Sec A Instructor	A_1	A_2…	A_7	$\mathbf{X_8}$	A_9	A_{10}	A_{11}	A_{12}	
Sec B Instructor	B_1	B_2…	B_7	B_8	B_9	B_{10}	$\mathbf{X_{11}}$	B_{12}	
Assessments In Common	$\mathbf{att_1}$	**D**		$\mathbf{Q_{RR}}$	$\mathbf{MT_{RR}}$		$\mathbf{Q_{LA}}$	$\mathbf{att_{12}}$	**FE**

Setting

The setting for our study is a research-focused university with a multi-section (11 instructors) Calculus 1 course primarily aimed at Business and Economics majors, though the course shares most core material with the science-oriented Calculus 1 courses at the same institution. Most students in the course this particular year were using "clickers" in at least one of their non-math courses in their first semester.

Interventions took place in Section A with 200 and Section B with 150 students (based on maximum enrolment). The tenured faculty instructors for the two sections (whom we will refer to as Instructors A and B) had strong teaching records in terms of length of experience, student evaluations and anecdotal department opinion. For this and the previous year, both instructors used "clicker" personal response devices to enhance classroom interactivity, asking 1- 2 such questions per hour on average. Otherwise, class time was primarily spent on relatively traditional lecture (concepts introduced at the blackboard, worked examples) with some directed whole-class discussion (one person speaking at a time, instructor leading the discussion). Both were receptive to student questions during class. This was the third consecutive year teaching this course for both instructors, and both had used "clickers" in the year prior to our study.

The intervention instructor for our study (Instructor X) was a graduate student who had taught this course once, as well as two other courses more recently. This instructor had used "clickers" in these latter courses, and participated in professional development activities promoting high engagement teaching methods. Though familiar with some recent results in education research, this individual would not qualify as a "science education researcher", characterized as expert in evidence-based teaching methods by a recent study of biology instructors (Andrews, Leonard, Colgrove, & Kalinowski, 2011) with the concern that existing positive results occur when such a researcher takes over as an instructor; this may distinguish our study further from other implementation studies of active learning pedagogy.

Assessments Prior to Intervention

A baseline of student abilities was established using three instruments, based on their predictive value for standard assessments in recent years:

- *Calculus diagnostic*: a 20-minute in-class test of prior calculus knowledge in the second week of term, with a mixture of "standard" procedure-based problems and conceptual problems.
- *Math Attitudes and Perceptions Survey*: online, measuring expert-like orientation to the discipline based on the CLASS Physics survey (Adams et al., 2006); student orientation to expert-like thinking about the subject is a known factor in course performance. Consists of 42 statements about perceptions, dispositions, study habits and beliefs about mathematics using a Likert scale, where students are scored based on their agreement with surveyed mathematicians.
- *Precalculus quiz*: online, problems chosen from a locally-developed multiple-choice placement exam; student scores on this subset were found in the previous year to predict final grades as well as high-school mathematics grades.

Though established instruments were considered, such as the Precalculus Concept Assessment (Carlson, Oehrtman, & Engelke, 2010) and the Calculus Concept Inventory (Epstein, 2006), they were ultimately not feasible to implement due to the restricted time available to assess the students.

Learning objectives

Prior to the intervention weeks, learning objectives for the topics were established by the entire research team to provide the objectives for instruction for those involved with the classroom. In addition to these stated objectives, instructors were aware of the homework problems and previous years' exam problems that students would be expected to solve as examples of assessment.

We claim that the two topics chosen, related rates and linear approximation, are relatively independent items in the course: the former draws on the notion of derivative as rate, implicit differentiation, and word problems with geometric objects, while the latter is more closely connected to the graphical interpretation of the derivative and estimation. Certainly, we expected student performance on assessment items to be influenced by their learning from earlier in the course, though in our future analysis we will attempt to account for this based on performance for diagnostic items and other topics on the final exam.

Assessments During the Intervention Weeks

The three instructors were responsible for teaching the agreed-upon material in the allotted times and allowing for 15-minutes topic-specific quizzes at the conclusion of the intervention weeks' class periods. The quiz format was a single, larger problem with intermediate prompts, developed in isolation from the instruction (by the third author) based on the learning objectives.

In the quiz of related rates, involving a growing conical pile of sand, students were asked as separate steps to draw and label a diagram, identify the rate requested in the problem statement, determine the rate relation, and finally solve for the requested rate at the specified moment (a standard problem may have this final step as the only prompt). These were intended to measure achievement of the stated learning objectives and also match up with the existing analysis of components of related rates problems, particularly the work of Martin (2000) which lists seven steps for solution of geometric related rates problems and reports the results of assessing students on specific steps.

The second topic quiz, which involved linear approximation of e^x, students were asked first to compute the value of the approximation using the tangent line at a specific point, then to decide if they had computed an overestimate or underestimate, then to draw the appropriate tangent line on a provided graph of the exponential function and label both the point of tangency and the approximating point, and finally to compute an estimate of the error involved. The tangent line drawing component, this item is rarely measured in local assessments, was inspired by poor student performance in observed group work in recent years as well as existing research about the relation of tangent lines to functions and their derivatives (Asiala et al., 1997).

Classroom Observation

To quantify the level of activity and time spent on various tasks, we employed the *Teaching Dimensions Observation Protocol* (TDOP) instrument (Hora & Ferrare, 2009), developed as part of an NSF-funded project at multiple institutions of higher education, where an in-class observer codes instructor behavior and (expected) cognitive demands upon the students in 5-minute intervals. This permitted a characterization of classroom activity of the control sections and experimental sections, summarized in Table 1 below. Observation was performed during the intervention weeks as well as a sample of classes outside of those weeks to confirm that the regular instructors were employing their usual instructional style during the experiment.

Assessments after the Intervention Weeks

The intervention instructor was excluded from the development of all of the study's assessment items, while the regular instructors were excluded from the development of the quizzes but involved in their usual capacity in the development of standard test items. These included two midterm tests: the later one featured an identical question about related rates and otherwise covered similar material. The final exam was in common with the other ten course sections, and all the regular course instructors were permitted input. As is typical for this course, the final exam prominently featured procedural computation along with a few conceptually-focused items.

The midterm test featured a common related rates problem: determining the rate of change of water depth of two tanks with the same height and volume, one a cylinder and the other an inverted cone, being filled at the same rate. The final exam featured another "tank" problem with a conical tank draining into a cylindrical one, thus requiring the connection of rates between the two tanks.

The intervention week for linear approximation occurred late in the term, so the final exam was the only common assessment after the quiz. The final exam question required students to compute a linear approximation for the natural logarithm, compute an error bound, discuss the error and later compare with the quadratic approximation at the same point.

Results of the Research

We have classified our results into three categories:

- Diagnostic items used to establish performance baselines for the two sections.
- Instructional methods comparison based on the classroom observation to quantify the different uses of class time.
- Student performance on the topic assessments.

Diagnostic Items

Our attitude and precalculus assessments indicated the student populations were similar to those of the previous year. On these and the new calculus diagnostic, the students in both sections achieved similar score distributions. We were not concerned about identical baselines due to the “crossover” structure used in the comparison, but such data establish these as typical sections in this course that are roughly comparable.

Methods Comparison

The Teaching Dimensions Observation Protocol (TDOP) codes classroom activity in 5-minute slices. For each instructor, Table 1 shows the average number of slices containing the described activity (slices can contain more than one type of activity). Class time during the intervention weeks was very similar for each of Instructors A and B, so the average over all observations is reported here. The categories are:

- **Admin**: Classroom announcements, hand out or pick up of paper, or other activity not related to content.
- **Lecture: new item**: Instructor presents new material/theory/ideas.
- **Lecture: example**: Instructor presents worked example.
- **Lecture: interactive**: Instructor leads classroom discussion by posing questions to students that receive responses (rhetorical questions not included here) and reacting to those responses.
- **Student Tasks**: Students are directed to work alone or in groups on a task,
- **Clicker Question**: Instructor poses an in-class voting question (multiple choice), students are given time to think and choose their response, possibly with peer discussion.
- **Q from Student**: Student asks a question to which the instructor responds.

The overall numbers from the observation protocol indicate that Instructors A and B both include significant class time leading class-wide discussions and responding to student questions, but the majority of time is spent on one-to-many lecture, with Instructor B showing a relative preference for drawing concepts from worked examples, compared to Instructor A’s more theoretical approach (usually followed by examples). We also see that Instructor X, the intervention instructor, does include lecture time, it is just in a minority (42% of the 5-minute blocks), while students are much more active. One interesting column of note is the “Question from Student” count: whereas both Instructors A and B have classroom cultures where students will ask unsolicited questions which are incorporated into the lecture as possible, the higher-engagement classroom of Instructor X had considerably fewer student questions during class time. We note that in the intervention classes, students were encouraged to ask questions in the pre-assignments submitted online, with the most noticeable concerns addressed in class (obviating the need to spend time in class) and further that the students were provided structured hand-outs, leading to a much more structured experience with some reference material.

Inst	5-min Slices	Admin	Lecture: theory	Lecture: example	Lecture: interactive	Student Tasks	Clicker Q	Q from Student
A	72	6.3%	29.5%	23.5%	18.4%	2.1%	3.6%	16.5%
B	109	6.6%	22.3%	36.3%	18.0%	0.0%	1.0%	15.9%
X	39	5.1%	20.6%	21.3%	20.1%	10.7%	16.9%	5.4%

Table 1: Proportion of 5-minute slices containing the described activity (slices can contain more than one type of activity) for each of the instructors.

Student Performance

Key results from our assessments are summarized in Tables 2 though 6. Tests of significance were performed for the proportions of students demonstrating a specific skill, either in a binary fashion (a row with its own p-value) or in a set of mutually exclusive categories (multiple rows with single p-value). We excluded the students who were not present for the instruction (who did not write a quiz) from our analysis; this was a considerable number for the second intervention week due to an external event attended by many students.

The data from our immediate (quiz) assessments support a positive answer for our first research question, and the follow-up (midterm) assessment for the related rates material supported a positive result for the second question.

In the related rates quiz, students in the experimental section were more likely to draw a sensible diagram and be able to start and complete a solution compared to the control section (Table 2). It is worth noting that the control section students spent most of their class time seeing extensive worked examples involving similar triangles and an inverted conical tank; this may have led many of them to an action conception of related rates problems, with a number of students in the control section drawing an *inverted* cone for the sand pile shape in the quiz, and more of them using the (efficient) proportional relation of radius and height before proceeding to the relevant time derivative. Students in the experimental section, where the cone geometry appeared in the reading but was not addressed in class, were less likely to make the correct proportional relation prior to computing the derivative, but were overall still more likely to succeed.

Section	**X**	**B**	**diff**	**χ^2 test**
Correct diagram	75%	60%	+14.2%	$p < 0.02$
Full solution	29%	24%	+5.0%	$p < 0.02$
Some useful work	59%	52%	+7.4%	
Blank or no significant work	11%	24%	-12.4%	
N	**177**	**131**		

Table 2: Related Rates Quiz. Report of students showing a sensibly-labeled diagram in each section, and a categorization of student performance in solving a problem involving a growing conical sand pile.

Results from the related rates midterm problem (Table 3) suggest that the control section students did not move past their action conception of related rates insofar as they used a proportional relation between radius and height for both the conical *and cylindrical* tanks, even though the radius is constant in the latter. Perhaps based on feedback from the midterm, it appears that students in both sections were able to sort out the correct radius dependence by the time of the final exam (Table 4).

Section	X	B	diff	χ^2 test
Correct radius dependence	59%	42%	+17.5%	p < 0.01
Changing radius for cylinder	26%	46%	-20.0%	
Constant radius for cone	7.5%	4.7%	+2.7%	
Blank or no significant work	6.9%	7.1%	-0.2%	
N	**174**	**127**		

Table 3: Midterm Related Rates Problem. Categorization of student performance in solving a paired problem with a conical and a cylindrical tank both

Section	X	B	diff	χ^2 test
Correct radius dependence	74%	64%	+10.0%	p > 0.2
Changing radius for cylinder	14%	16%	-2.5%	
Constant radius for cone	6.3%	12%	+5.3%	
Blank	6.3%	8.5%	-2.2%	
N	**174**	**129**		

Table 4: Final Exam Related Rates Problem. Categorization of student performance in solving a joint problem with a conical tank draining into a

With the control and experimental sections reversed for the linear approximation material, we see apparently mixed results in the immediate quiz (Table 5); in this case students in the control section had seen correct pictures of linear approximation drawn and the idea of an underestimate when the tangent is below the function repeated multiple times during lecture. Students in the experimental section did not see or discuss the over- versus underestimate issue as much, and we see this in the split in the second row of Table 5. A dramatic result, however, is seen in the different attempts at drawing the correct tangent line; it appears that seeing the correct picture in a few cases was not sufficient instruction for the majority of the control to be able to draw the correct tangent line on the provided function graph. The experimental section students had read and been prompted to interact with an online applet in

Section	X	A	diff	χ^2 test
Compute approximation	52%	60%	-8.3%	p > 0.2
Correctly choose underestimate	66%	82%	-15.9%	p < 0.01
Draw tangent and label both points	42%	21%	+20.5%	p < 0.0001
Correct tangent, zero or one labels	24%	26%	-2.0%	
Tangent at approximation "x"	3.7%	29%	-24.8%	
Tangent elsewhere or not tangent	27%	19%	+8.6%	
No drawing	3.8%	6.0%	-2.2%	
N	**106**	**133**		

preparation for class, as well as having time in class to attempt their own sketches with instructor follow-up for comparison.

Results from the final exam's linear approximation problem are not conclusive (Table 6), though the experimental section was more capable in using the second derivative in analysis of the error. This is encouraging from an instructional standpoint, as this is often considered the most difficult aspect of linear approximation by the local instructors and students.

Section	X	A	diff	χ^2 test
Compute approximation 1	72%	78%	-5.6%	$p > 0.2$
Compute approximation 2	87%	91%	-4.2%	$p > 0.2$
Use of 2nd derivative in error bound	81%	66%	+15.0%	$p < 0.02$
Error analysis: any useful work	92%	85%	+7.5%	$p > 0.2$
N	**105**	**132**		

Table 6: Final Exam Linear Approximation Problems. Student performance on items related to linear approximation and its error on the final exam.

In summary, we observed better performance on conceptual parts of the related rates assessments (more of the students demonstrated an Action or Process understanding of various concepts), and a larger number of students able to demonstrate the correct picture for linear approximation (66% versus 47% of the class could draw the correct tangent line, while 42% versus 21% could do so and label the relevant points involved in approximation), for the higher engagement section in each case. Performance in both sections was very close on computational items and concepts more strongly tied to earlier parts of the course. The data from the final exam, which used relatively "standard" problems and thus had less fidelity in exposing specific concepts, was not especially supportive of our second research question, though students in the experimental section were more likely to connect the second derivative to the computation of an error bound in linear approximation.

Remaining Project Work

At the time of writing, data collection and analysis is ongoing, and a department teaching assistant has also been employed to assist in further data collection from the completed assessment items.

Plans for further work include:

- Matching up attitude data with performance data; are students who report certain dispositions more or less likely to learn under the different teaching conditions?
- Formally establishing inter-rater reliability for the assessment coding.
- Tracking student learning through the term by linking them on assessment items (this report compares the populations separately item-by-item); this may also permit adjustment based on student learning in other topics, such as ability to compute derivatives.

Acknowledgements

The authors would like to thank Matthew Thomas (University of Arizona) and the conference referees for their extremely valuable feedback on the initial proposal.

References

Adams, W., Perkins, K., Podolefsky, N., Dubson, M., Finkelstein, N., & Wieman, C. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, *2*(1). doi:10.1103/PhysRevSTPER.2.010101

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 education*, *10*(4), 394-405. doi:10.1187/cbe.11-07-0061

Asiala, M., Cottrill, J. F., Dubinsky, E., & Schwingendorf, K. E. (1997). The development of students' graphical understanding of the derivative. *The Journal of Mathematical Behavior*, *16*(4), 399-431. doi:10.1016/S0732-3123(97)90015-8

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2003). *How People Learn: Brain, Mind, Experience, and School*. *Psychology* (Expanded.). Washington D.C.: National Academy Press.

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*, *33*(5), 352-378. Retrieved from http://www.jstor.org/stable/4149958

Carlson, M., Oehrtman, M., & Engelke, N. (2010). The Precalculus Concept Assessment: A Tool for Assessing Students' Reasoning Abilities and Understandings. *Cognition and Instruction*, *28*(2), 113-145. doi:10.1080/07370001003676587

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, *69*, 970. doi:10.1119/1.1374249

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science (New York, N.Y.)*, *332*(6031), 862-4. doi:10.1126/science.1201783

Dubinsky, E. D., & 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* (pp. 275-282). Springer Netherlands. doi:10.1007/0-306-47231-7_25

Engelke, N. (2007). *Students' understanding of related rates problems in calculus*. Arizona State University.

Epstein, J. (2006). The Calculus Concept Inventory. In D. Deeds & B. Callen (Eds.), *Proceedings of the National STEM Assessment Conference* (pp. 60-67). Retrieved from http://www.openwatermedia.com/downloads/STEM(for-posting).pdf

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.

Hora, M., & Ferrare, J. (2009). *Structured observation protocol for instruction in Institutions of Higher Education (IHEs)*. Madison, WI.

Martin, T. (2000). Calculus students' ability to solve geometric related-rates problems. *Mathematics Education Research Journal*, *12*(2), 74-91. doi:10.1007/BF03217077

Michael, J. (2006). Where's the evidence that active learning works ? *Advances in Physiology Education*, *30*, 159-167. doi:10.1152/advan.00053.2006.

Speer, N. M., Smith III, J. P., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99-114. doi:10.1016/j.jmathb.2010.02.001

Tziritas, M. (2011). *APOS Theory as a Framework to Study the Conceptual Stages of Related Rates Problems*. *Analysis*. Concordia University.

Weller, K., Clark, J., Dubinsky, E., Loch, S., McDonald, M., & Merkovsky, R. (2003). Student performance and attitudes in courses based on APOS Theory and the ACE Teaching Cycle. In A. Selden, E. Dubinsky, G. Harel, & F. Hitt (Eds.), *Research in Collegiate Mathematics Education V* (pp. 97-131). Providence: American Mathematical Society.

TOWARD A DESCRIPTION OF SYMBOL SENSE IN STATISTICS

Samuel A. Cook & Tim Fukawa-Connelly
Wheelock College & The University of New Hampshire

This is a first attempt to describe how students might develop a statistical symbol sense and what such a symbol sense entails. The paper first presents a genetic decomposition for a symbolic understanding of the arithmetic mean, the standard deviation and the standard error of the sample means of a sampling distribution by drawing on Sfard's (1991) process-object duality. There is currently little research in which to ground a genetic decomposition and, as a result, the one presented here draws primarily upon the authors' experience teaching statistics. It needs extensive testing and revision, but it is meant to serve as a starting point for future investigations into students' development of understanding of statistical symbols. The paper ends by describing some important attributes of a symbol sense in statistics based upon Sfard's (1991) framework and Arcavi's (1994) description of a symbol sense in mathematics.

Keywords: statistics, symbol sense, process-object, genetic decomposition

Introduction and Motivation.

Arcavi's (1994) paper on symbol sense in mathematics, while not explicitly situated within the tradition of semiotics adopted the position that symbolic understanding and fluency was an important component in knowing and doing algebra. That is, fluency with particular types of mathematics required fluency with a broad range of presentations, including the symbolic. In particular, Arcavi claimed that students should, at minimum.

- Know how and when symbols can and should be used in order to display relationships
- Have a feeling for when to abandon symbols in favor of other approaches.
- Have an ability to select a representation and, if necessary, change it.
- Understand "the constant need to check symbol meanings while solving a problem, and to compare and contrast those meanings with one's one intuitions or with the expected outcomes of that problem" (p. 31).

It is certainly true that algebraic skills do support students' ability to do and understand statistical concepts (Lunsford & Poppin, 2011). As a result, we argue that there are reasonable analogues to Arcavi's habits and skills in the realm of probability and statistics that are important to consider. While there have been investigations of students' understanding of measures of center (Mayen, Diaz, Batanero, 2009; Watier, Lamontagne, & Chartier, 2011), variation (Peters, 2011; Watson, 2009; Zieffler & Garfield, 2009), and students' preconceptions of the terms related to statistics (Kaplan, Fisher, & Rogness, 2009). A literature search of the titles, keywords and abstracts of all papers in the Journal of Statistics Education and the Statistics Education Research Journal suggest that none had a primary focus on investigating and exploring students' use and understanding of the symbolic system of statistics; although, one paper did draw upon the onto-semiotic tradition to describe student errors related to representations of the mean and median (Mayen, Diaz, Batanero, 2009).

Theoretical Perspective.

The arbitrary nature of symbolic representations

Symbolic representations of mathematical ideas are regarded as particularly critical due to Hewitt's (1999, 2001a, 2001b) distinction between arbitrary and necessary elements of the mathematical system. Hewitt notes that names, symbols and other aspects of a representation system are culturally agreed upon conventions. For those who already understand them they might feel sensible, but "names and labels can feel arbitrary for students, in the sense that there does not appear to be any reason why something has to be called that particular name. Indeed, there is no reason why something has to be given a particular name" (1999, p. 3). Hewitt continues by differentiating between those aspects of a concept used by a community of practice which can only be learned by being told and then memorizing, which he labels arbitrary, and those which can be learned or understood through exploration and practice, which he labels necessary. Additionally he notes that for students to become proficient at communicating with established members of the community of practice, they must both memorize the arbitrary elements and correctly associate them with appropriate understandings of the necessary elements.

This problematizing of symbolic representations and the linkages between symbolic representations and the concepts being symbolized is at the heart of the field of semiotics. Eco (1976) gave the term semiotic function to describe the dependence between a text and its components and between the components. The semiotic function relates the antecedent (that which is being signified) and the consequent sign (or that which symbolizes the antecedent) (Noth, 1995). When considering the statistical community and the representation system in use within that community, there is a defined complex web of semiotic functions and shared concepts that "take into account the essentially relational nature of mathematics and generalize the notion of representation: the role of representation is not totally undertaken by language (oral, written, gestures, …)" (Font, Godino, & D'Amore, 2007, p. 4). Throughout this paper, we recognize the inherent arbitrary nature of much of the symbolic system of statistics and draw on the notion of semiotic function as a means of linking a particular representation with the relevant concept. In doing so, we articulate specific linkages that students should be developing and describe some of the difficulties and potential pitfalls of the symbolic system.

Sfard's Process-Object Duality

In her seminal work on the dual nature of mathematical concepts Sfard aligned herself with the semiotic tradition by noting that mathematical objects are inaccessible to our senses and "the sign on the paper is but one among many possible *representations* of some abstract entity, which by itself can be neither seen nor touched" (1991, p. 3). We set aside the ontological question about whether mathematical objects have an existence in a Platonic Realm and concentrate on the psychological questions involved. Following Sfard, we give two different descriptions of ways to understand a mathematical object.

The first means of understanding a concept, which Sfard describes as less abstract, less integrated, and more detailed than the second, is as a *process*. In holding this 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 process view of a concept is one that includes change and one that requires sequential steps and actions.

The second is that an individual might see a mathematical entity as an *object* and is therefore capable of referring to it as if it is 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. Sfard claims that this type of understanding is aligned with the type of structural thinking common in modern mathematics, drawing upon abstract definitions and theorems and should be understood as more advanced. A student who holds an object-type of conception understands the concept as a member of a category (not using the mathematical definition of a category here) about which questions can be asked and answered, including explorations of the general properties of that category of concepts, and various relations between the concepts in the category.

The dual nature of the mathematical concepts comes through in verbal, visual and symbolic representations. While an individual might perceive a concept in a certain way, Sfard asserts that certain representations "appear to be more susceptible of structural interpretation than others" (p. 5). She asserts that those representations which are more compact, which can be completely manipulated, and which can help make abstract ideas tangible better support an *object* interpretation while those that cannot, and especially verbal descriptions, better support a *process* interpretation.

Sfard describes three stages in a student's development from a process-oriented to an object-oriented understanding. She calls them *interiorization, condensation*, and *reification*. During the interiorization stage the student is becoming familiar with a process and performing it on a lower-level object. Through the process of interiorization students are to become familiar enough with a process that they can carry it out mentally without needing to actually perform it, and this interiorization would then allow a student to analyze the process.

The second phase is that of *condensation*, which is a phase where the student becomes capable of seeing and thinking about a process as a whole. Sfard wrote, "this is the point at which a new concept is "officially born" (1991, p. 19). She continued by noting that during this phase, the student becomes increasingly able to switch between different representations of a concept. Both interiorization and condensation should be understood as gradual and quantitative shifts in a student's understanding. The third phase, *reification*, is both an immediate and qualitative shift. It happens at the point where a student is able to apprehend the concept as an *object*.

A Notion of Symbol Sense

Many of the habits and skills that Arcavi (1994) described have a natural analog in statistics. Most important of these is knowing how and when symbols can and should be used. In mathematics a symbol typically represents an unknown or is defined to represent a single mathematical concept; however, in statistics, symbols often carry multiple layers of meaning. For example, both $\bar{x}$ and μ are well defined as an arithmetic mean; however, each has a second layer definition defining what type of data set the arithmetic mean comes from; $\bar{x}$ is the mean of a set of sample data and μ is the mean of a set of population data. This additional layer of information is crucial in displaying relationships. That is, the student should encode the semiotic function linking the representation (symbol) and the concept of mean, and should be a part of a student's statistical skill set at the end of a course.

Arcavi also recommends knowing when to abandon symbols in favor of other approaches. This has a non-mathematical application to statistics. While statistical procedure revolves around the relationship between symbols and their relationship to a sample and population, the practical use of statistics is much less technical. In many instances statistics is the tool used to explain or reason about something in a different discipline such as psychology or biology; disciplines that are not necessarily rooted in mathematics. It is important to be able to abandon descriptive symbols in favor of concise statements such that a hypothesis or a conclusion can be interpreted without understanding what a symbol represents. A student should not only be able to abandon formal symbol representation, but be able to “translate” symbolic statements into something easily understood by all.

Finally, Arcavi states that a constant check of symbol meanings during problem solving is needed. In statistics, the multi-layered meaning of symbols makes this important. Additionally, there are general mathematical symbols that are mathematical operators; however, in statistics it is a general rule that a Greek symbol represents a population summary and a Roman symbol represents a sample summary, but there are times when Greek and Roman symbols are nearly indistinguishable such as with Nu. A student might see N = 25, and not understand why one is to use capital N for a population and lower-case N for a sample while a statistician might be surprised that the student does not recognize Nu! Thus, from the different perspectives, a symbol might be completely reasonable or seemingly arbitrary. This continues with inclusion of symbols such as “$\sum$” as operators, rather than conveying information about a population, will sometimes confuse students and makes these general rules less clear than intended.

An Expansion of Arcavi’s List

The following section will briefly outline a few ideas that might be understood as forming part of a statistical symbol sense. It is important that students have a clear understanding of relevant terms and be able to correctly associate each term with the most appropriate symbol. Beyond that, students should:

- Understand, in the context of a given problem, which symbols represent constants (even if unknowable) and which represent values that can vary.
- Understand that symbols which are constant for a given problem can also be understood as varying across problem contexts.
- Possess a feeling for when symbols should be used to display relationships and when visual representations better convey appropriate information.
- Demonstrate an ability to read symbolic expressions for meaning, both in the context of the problem, while also connecting them to their abstracted concepts.
- Consistently check the meaning of the symbols against the problem and with their own intuition.
- Possess an understanding of the difference between different symbols that represent the same basic concept (such as a sample mean versus a population mean).

In service of that, we'll relate the three vertices of the triangle below (verbal/symbolic/visual) and explain how we believe the students are likely to process through their symbolic understanding.

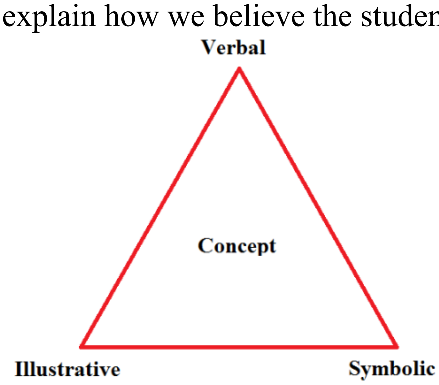


Diagram 1: A visual representation of the relationships between a concept and possible forms of representation of that concept

Research Aims.

This theoretical report aligns itself with Arcavi's (1994) work and the tradition of semiotic research and is situated in the context of statistics education. We will draw upon Sfard's (1991) notions of process and object as well as the psychological process of interiorization, condensation, and reification in order to:

- Create a hypothesized learning trajectory of students' development of a symbol sense in statistics.
- Illustrate this learning trajectory with common examples from an introductory statistics course
- Describe future research needed to better establish the validity of our proposed learning trajectory

Due to the dearth of research in the field, our trajectory will be based primarily upon the understandings of our students' learning that we have developed through our instruction and one small-scale research study (Kim, Fukawa-Connelly, & Cook, 2012). We align ourselves with Contrill, et al. (1996) in terms of describing our goals and how we understand their value. Essentially, this hypothetical learning trajectory will likely be continually evolving, and we …

> Do not suggest that it is a "true" description of what is going on in the minds of the students we are observing, nor that it is in any sense "proved' or even established. We only claim that it is a tentative description that has the following value:
>
> - It provides a method for making sense out of a large amount of qualitative data.
> - It provides a language for talking about the nature of learning particular topics in mathematics.
> - It has the potential to suggest pedagogical strategies that could improve the extent to which this learning takes place.
>
> The first two points are matters of judgment, but the third can be evaluated and it is part of our paradigm that the instruction based on our genetic decompositions eventually be evaluated to see what effect it has on student learning. (Cottrill, et. al., 1996, p. 170)

In short, the proposed learning trajectory will be a theoretical model that makes testable predictions about students' abilities and the growth of those abilities in statistics. The long-term goal of this work is to guide both research and curriculum design efforts around students' understanding of the symbolic system of statistics.

A learning trajectory for statistical symbols and symbol sense.

With the expansion of statistical software and technology the need for mathematical computation is decreased and less emphasized (Garfield & Ben-Zvi, 2008). Without computation, a student's ability to understand a symbol and have a symbol sense is more difficult. Understanding statistical symbols as objects as opposed to computational components and knowing when to identify a symbol as each can be shown in the trajectory from process to object that includes the stages of interioization, condensation and reification. Mayen, Diaz, and Batanero (2009) have reported that students experience difficulty with the symbols related to sample and population mean despite finding that the students had an ability to distinguish them, thus suggesting that for students, relating the concepts to their symbols is difficult. Moreover, there is research suggesting that students' prior understandings of terms used in statistics colors how they make sense of the statistical concepts (Kaplan, Fisher, & Rogness, 2009). When taken together, this research suggests that symbol sense is difficult for students and that students will bring their previous symbolic and verbal understandings, sometimes developed in mathematics classes sometimes in daily life, to their statistics learning in ways that may both support and/or hinder their statistical learning.

The Mean

A process orientation of the mean

We assert that students begin to understand the mean as an entirely symbolic-verbal process. They are to "add up the values and divide by how many they have." Symbolically, that would look like either an iterative process of $X_{next} = X_{now} + x_i$, a process of taking an extant number and deriving the next sum by adding a value and repeating that process until all numbers to be included in the sum have been added. At the undergraduate level, we claim that due to their experience with addition, they might be able to apprehend the entire summation, although typically without the sigma notation. As a result, they hold a process conception where "add all the values" is one step in the process and "divide by the number of value that you have" is the next step. Symbolically, that might look like:

1)

2) $X_{next} = X_{now} / n$

We assert that even at this point, students have a semantic function linking '*n*' and *the number of members*. In short, at this point, they see both μ and $\bar{x}$ as a command to carry out the same process (as described above) and therefore, cannot apprehend a meaningful difference between them. As a result, when students have a process-conception of the mean, if they correctly associate μ and $\bar{x}$ with populations and samples respectively, it is through a memorized distinction of the sort, "Greek letters correspond to populations."

Interioization of Mean

Determining the mean only relies upon arithmetic operations that students are generally comfortable with. The arithmetic process of determining a mean is one that is quickly taught and understood. It is at this point that students start describing the process verbally, creating a linkage between the symbolic and verbal descriptions. Before identifying that there is a difference between μ and $\bar{x}$ a student will describe the mean as the additive center of data. When thinking of μ and $\bar{x}$ a student who has interiorized the process might look for a difference in the *process* rather than in the set of values that the measure is describing (Kim, Fukawa-Connelly, & Cook, 2011) . For example, a student might describe the difference between $\bar{x}$ and μ as "add up and divide" vs. "put them in order, add up and divide." As students start to condense their understanding they will increasingly be able to note that the process for determining the means might not be different but rather that they draw on different sets of data, and, as a result need a different symbolic designation.

Condensation of the Mean

When the student is familiar enough with the process to understand it, without relying on the arithmetic process, the student will be capable of seeing the mean as a whole and not as the result of a formula. The capability of seeing a mean as the arithmetic center of a group of numbers allows the student to focus on the group of numbers and not on the process of finding the mean. It is during this stage that students can describe what will approximately happen to a mean if a particular data point is changed or another point is added to the data set. When the focus is the data and not the mean, the difference between μ and $\bar{x}$ begins to present itself. It is at this stage that a student begins to develop a symbol sense for the mean. If the group of numbers represents every member a population then the mean is μ and if the group of numbers represent a subset of the population it is $\bar{x}$. It is at this point that a student might be able to create a linkage between the symbolic representation and a visual representation showing the mean as a balance point when the data points are arrayed on a number line. Similarly, it is at this stage that the students begin to recognize $\bar{x} = \frac{\sum_1^n x_i}{n}$ as an equation rather than formula or rule for calculating the value of the mean meaning that they recognize that one side of the equation can be substituted in place of the other without changing the truth-value of or statistical claims in a sentence.

Reification of the Mean

Once a student is able to focus on the data and the whole process associated with that data the student is able to apprehend the mean as an object and finally make a connection between a visual representation and the symbolic representation. The student can now see the mean as something that exists on its own and not because of an arithmetic process. A clear understanding of what μ and $\bar{x}$ represent becomes present and a student will be able to relate the two. Understanding the mean as an object allows one to relate $\bar{x}$ to its corresponding μ because a student will be able to focus on the data and not the process. As data from the population is added to data that $\bar{x}$ represents, a student will have an understanding of what is likely to happen to $\bar{x}$ and can see μ as a constant and $\bar{x}$ as variable. It is at this point that a student might apprehend an important distinction between μ and $\bar{x}$. In particular, a student might see μ as the additive center, and might think of $\bar{x}$ as a number close to μ, and not as much as the additive center, but rather as the additive center of a subset of the data.

The Standard Deviation

Developing a symbol sense of the standard deviation requires that students be able to operate with the mean, which requires, at the least, that they have experienced a condensation of the mean. In short, a student needs to be able to see the mean as something that can be operated on.

A Process Orientation of the Standard Deviation

Holding a process orientation to the standard deviation means that students make sense of it by actually calculating it with a focus on the 'calculating' side of the formula: $\sqrt{\frac{\sum\left(x_i - \bar{x}\right)^2}{n-1}}$ (or division by n) while attempting to memorize the symbols s and σ. Yet, in a reform class, given the heavy use of technology, students are unlikely to perform this calculation at all, and at most once. As a result, we argue that students are expected to skip this stage in the developmental trajectory or pass through it based upon their understanding of symbols from their mathematics classes. At best, they are expected to develop the symbolic understanding by linking it to the verbal description; subtract the mean from each data point, square the difference, sum the squares, divide, and then square root. Yet, even being able to describe the process suggests that the students are progressing to the interiorization stage. At this point, the students have no connection between the symbolic and the visual representations for the standard deviation. Moreover, while a student might be able to describe the concept, "a measure of the variation," this description is likely to be wholly disconnected from any symbolic representation of the standard deviation.

Interiorization of the Standard Deviation

At this stage in their development of symbolic understanding of the standard deviation, students can develop a linkage between the verbal and symbolic representations. In particular, students can describe what each of the pieces of the symbol does in terms of the process described above and think through each of the actions in calculating a standard deviation for a particular data set. A student might ask why a particular aspect of the formula was chosen, for example, why does the process require squaring and then square rooting instead of using an absolute value or why does the statistic include division by (n-1) instead of division by n. Yet, at this point, the most mathematically and statistically meaningful responses to this question are likely beyond the students' verbal developmental stage. Thus, we typically are left with an unsatisfying response about underestimation of variation in the population that we correct by dividing by a smaller number.

Condensation of the Standard Deviation

At this stage in their development of symbolic understanding of the standard deviation the student can develop a further linkage between the verbal and symbolic representations. In particular, the student can link the verbal description or "a measure of the variation in the data points" with the symbolic representation of the calculating side of the formula. That is, the student can explain how the calculating formula captures a measure of variation. Moreover, the student is able to think about and analyze the process of calculating the standard deviation as a whole, meaning that at this stage they are able to think about how changes to the data set would give rise to approximate changes to the standard deviation in terms of direction and magnitude of change. For example, a student might be able to say, "replacing that data point with x would result in a small decrease in the standard deviation." It is also at this stage that the student starts to be able to apprehend the formula $s = \sqrt{\frac{\sum \left(x_i - \bar{x}\right)^2}{n-1}}$ as an equation or statement of equivalence, not as a rule for calculations. Similarly, at this point, the student no longer sees the right side, as a discrete collection of symbols that must be decoded individually. Rather, they are able to apprehend them together and see them as describing a process that can be either on-going or completed. That is, a student would recognize that one side of the equation might be substituted for the other without changing the truth-value of or statistical claims in the statement. At this point the students would also be capable of creating a link to a visual sum of squares type representation of the standard deviation and the calculating formula. While there is not typically a number-line representation given, one could be created that might now hold meaning for the students in terms of the calculating formula.

Reification of the Standard Deviation

Once a student is able to focus on the standard deviation as a representation of the variation of data the student is able to apprehend the standard deviation as an object. When seen as an object, σ holds distribution information that allows a student to visualize the outcome of the most common data points in a data set and make a connection between a visual representation and the symbolic representation. The student can now see the standard deviation as something that exists on its own and not because of an arithmetic process. A clear understanding of what σ and s represent becomes present and a student will be able to relate the two. This understanding allows the student to apprehend the standard deviation as an object that can be operated on. Once the standard deviation has been reified, a student will be able to understand the standard deviation of a data set and make sense of it without referring to a visual or verbal description. For example, at this point, a student might be able to answer a question that asks them to relate the size of the mean and the standard deviation to the skew of the distribution. At this point, the connection to the normal distribution graph that has standard deviations marked out might allow for meaningful connections to the symbolic form. In terms of a linkage to a verbal representation, it is at this point the relationship to Chebychev's theorem carries meaning. In particular, students can have a meaningful symbolic sense that no matter the distribution the proportion of data between 2 standard deviations is at least 75%, or, more generally, at least $\left(1-\frac{1}{k^2}\right)\%$ of the data is within k distributions of the mean.

The Standard Error of a Sampling Distribution

The standard error of a sample mean requires the creation of a sampling distribution; therefore, it would be helpful if students had a dynamic image in their heads of samples being created from the original population, each sample being of size n. Then, for each sample, the sample mean is computed and the distribution is created. To conceptualize this, understanding the mean and the standard deviation as objects is needed. The mean of the sampling distribution, the mean of all possible sample means, is the same as the mean of the original population. The standard deviation of the sample means is a measure of the spread of the sample means of size n from their mean. That is, this formula is meant as a measure of how spread out a population of sample means is. In order to make sense of this formula, it requires the student to have constructed a mental landscape with the ability to operate on at least two levels of abstraction; one is relatively low and is the original distribution, while the second is relatively high and asks students to contemplate the distribution of all possible sample means of size n where the individual samples are drawn from the original distribution.

A Process Orientation of the Standard Error of a Sampling Distribution

When considering the standard error, students are confronted with the simple looking formula: $\sigma_{\bar{x}} = \sigma/\sqrt{n}$. This formula requires an easy calculation (that we often make easier by picking 'friendly' sample sizes), and, has a relative simple-seeming explanation, "the standard deviation of the sample means." When students confront the equation $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, one of their first realizations should be the formula mixes notation for populations and samples (σ and *n and* $\bar{x}$, respectively). As a result, students need to have a decision rule that allows them to understand what is being described; is this formula describing a sample? A population? Moreover, it requires students to be able to understand when it is appropriate and valuable to mix elements from, what they have possibly understood as separate representational systems; those that represent data derived from a sample and those that represent data derived from a population.

In fact, this formula, and, in particular, the symbol $\sigma_{\bar{x}}$ is describing a moment in an entirely *new* distribution, one that is distinct from the original population, demands consideration of a sample of size n, is based upon the old distribution and requires the student to understand σ as an object. In this case, there is representation, $\sigma_{\bar{x}}$, that draws on the agreed-upon symbols for population standard deviation and sample mean to communicate this information.

When considering the representation on the right, $\dfrac{\sigma}{\sqrt{n}}$, students are given a formula that implies they should perform a calculation, again mixes elements from possibly separate representational systems, and, most importantly, could be understood as conveying an entirely different meaning of the standard error of the sample mean. In this case, students are confronted with four different possible representations of the same object (two verbal, two symbolic), to say nothing of the attempted graphical representations.

Interiorization of the Standard Error of a Sampling Distribution

Determining the standard error is the result of a simple arithmetic calculation; however, for a student to begin to understand the standard error a reification of the mean and standard error must be achieved. The student should look at the equation and read in terms of how the standard deviation of the sample mean compares with the original standard deviation. The student should ask themselves what division by the square root of *n* does, especially as *n* varies and identify what happens when *n* is 1. Students should understand that this would recapitulate the original distribution, both because each 'sample' would be exactly one individual (meaning that each individual in the population is then in exactly one sample) and because the symbols show that the square root of 1 is 1, and then the standard deviation of the sample means is the same as the standard deviation of the population because of division by 1. Then, the student should be able to explain how the value of the standard deviation of the sample means will change as the sample size increases by noting that σ is a constant and, then, division by an increasing value will cause a corresponding decrease in the final result. The students should imagine the distribution (the graphical representative) collapsing about the mean in a dynamic way [See Diagrams 1a and 1b for an illustration of this].

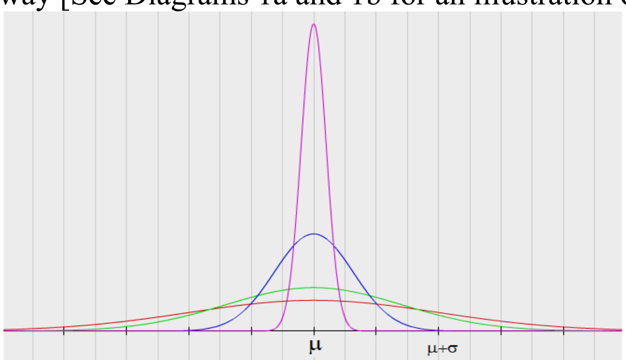


Diagram 2a: A normal distribution and the distribution of sample means from samples of size 2, 10 and 100.

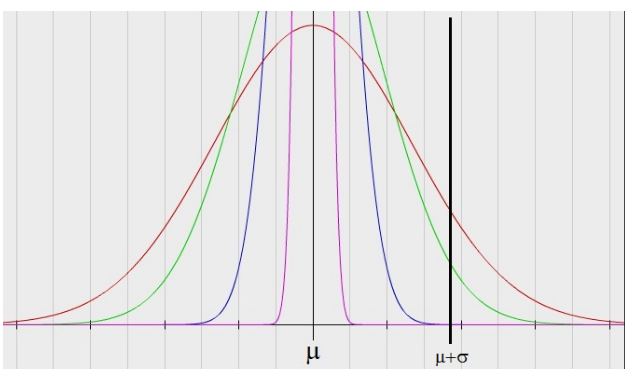


Diagram 2b: A normal distribution and the distribution of sample means (n = 2, 10 and 100) scaled towards the parent distribution

Condensation of the Standard Error of a Sampling Distribution

At this stage in their process of developing symbolic understanding of the standard error a student can describe a further linkage between the verbal and symbolic representations as well as their understanding of the relationship between the underlying distribution and the sampling distribution. At this stage, students will have encapsulated the process of generating a collection of sample means by repeatedly choosing elements from the underlying distribution and calculating the mean and standard deviation of those sample means. That is, they will clearly see a collection of sample means as objects that can be acted upon. But, during the condensation stage, they will also start extending that understanding to include the construction of "all possible sample means" of some particular size. This understanding can be concurrently visual, verbal and symbolic, although we rarely present a symbolic form. As a result, we argue that there is a missing link in the students' typical symbolic development of understanding of a sampling distribution.

In terms of linking verbal and symbolic representations, the student can link $\sigma_{\bar{x}}$ to a visual illustration of a data set that is made of many individual sample summaries of $\bar{x}$ (even if it is only a mental illustration). A student will be able to identify a transformation in the variation of the data by a product of $1/n$ and thus a transformation in the standard deviation of the new distribution of 1/root(n). A student who has reified the standard deviation will be able to see σ as an object that is constant and exists in a population and begin to see $\sigma_{\bar{x}}$ as a similar object that depends on a transformation that is dependent on the sample size. At this point the student will be able to understand that $\sigma_{\bar{x}}$ and $\sigma/\sqrt{n}$ are representative of the same idea and can be interchanged in a distribution summary without loss of meaning. A student who has reified the standard deviation will also have a dynamic idea that a minimum of 75% of the new data set will be within 2 of the transformed standard deviations. A demonstration of this, allows the student to see that, not only is the standard deviation of the new data dynamically shrinking, but it is also grouping around the population mean, mu. The knowledge that at least 75% of data is within the mean, allows the student to identify the condensed area that the new data takes up.

Reification of the Standard Error of a Sampling Distribution

A large amount of conceptual understanding of symbols is required for a student to begin to understand a standard error, including a vital need for a student to understand both a mean and standard deviation as an object and not only the result of a calculation. In the reification stage, the students gain the ability to make sense of the standard error of the sample means as a whole, both in terms of the formula and as a measure. In terms of how students apprehend the formula, the students no longer focus on individual elements of the formula, but rather they look at the formula as a statement of equality that allows substitution of either form in place of the other and they are able to look at the formula as a unit and operate upon it. Once a student is able to focus on the standard error as the representation of the variation of the sample means from a parent distribution, the student is able to apprehend the standard error as an object. When seen as an object, $\sigma_{\bar{x}}$ holds information about both the new distribution and the parent distribution; information that allows a student to visualize the likelihood of sampling events and in the case of an unknown parent distribution gives the student an inferential tool that will later be included in more advanced topics. The student can now see the standard error as something that exists on its own for each sample size and not because of an arithmetic process. In the section that follows we describe some of the reasons that symbolic understanding, and especially of the formula $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ may be problematic for students, and, how a statistical symbol sense would support their development of proficiency in using statistical tools.

Toward a Symbol Sense for Statistics.

Reading symbols for meaning related to the problem.

A student must be able to answer "What can vary?" and "What's constant, even if unknown?" to fully understand a problem. In the context of the formula for the standard error of a sampling distribution, students should be asking themselves these questions. Yet, the answers require a non-trivial ability to negotiate between contextualized and generalized understandings. At the most general, both σ and n can be understood as varying, the formula is applicable to all distributions, and, therefore, any σ. But, in most situations that the students encounter, they should be thinking in terms of a specific underlying distribution, which means that σ is fixed; although, it may be unknown (which the students should be able to discern). Yet, we want the students to understand that once the population, and thereby σ is fixed, that by changing sample sizes they create a large number of different sampling distributions. That requires students to understand the sample size n as able to vary and we should teach them to think this way. In terms of the level of understanding necessary to hold these types of understandings, we believe that students must have reified the concepts and symbolic understandings of mean, standard deviation and sample size and must be in the condensation phase, at minimum, for the standard error of the sample means.

To liken this to an element of algebra, when students consider quadratic functions, they should understand that $f(x) = ax^2$ gives rise to a quadratic, and, that for a particular instance, a is fixed, but we also want them to understand that a can vary and what that variation does to the function. Yet, they also need to be able to proceed into further contextualized problems where n has also been fixed and they, then, need to be able to picture the shape of the distribution and describe what effect n has on the shape of the distribution. Students might do this by drawing an appropriate picture of the distribution with ranges variation, as described by differences from the mean, marked. Students can make the transfer from their algebraic or functional understandings to the context of statistics (Kim, Fukawa-Connelly, & Cook, 2012) and recognize that certain elements of formulas work as translations or scaling factors. Yet, even when doing so, students have not necessarily developed a statistical symbol sense due to the need to reconceptualize what a variable is and how it might behave, as well as to be able to see the mean as an example of a possible variable. That is, while students might be able to rely on their mathematical understandings in order to compensate for weak statistical symbolic sense, to meaningfully relate the statistical symbols and formula to the visual or verbal displays likely requires that the students have an object-understanding of standard deviation, mean and sample size and have begun the condensation phase of understanding the standard error of the sample means.

The example of the standard error of a sample mean is an example of a concept that, when understood, makes understanding expected results straightforward. It is this concept of what is expected that is a building block of statistical inference. Students often dive into inference without conceptual understanding of what "should" happen under the premises provided. The ability to read expressions for meaning is a skill we should expect of statistics students. If a student has information about σ, then that student should have the ability to infer what outcomes for the sample mean are most common, and how they vary. This skill directly leads to the concept of "unlikely events" and a student can then infer what is likely versus what is unlikely by only understanding what the premise of the problem.

On Visualization and Selection of the Display

One of the challenges for students in understanding the sampling distribution is making sense of what individuals represent. They typically begin a statistics class by exploring data where an individual is a single measurement from one member of the population under study. This might be a heartbeat, count of siblings, or Likert scale rating, but, each number could be understood as describing one individual and often a person. That is, a single thing that could be visualized. When students start to consider a sampling distribution, the individual members of the population are now samples, and the measurement of each individual that we are considering is a mean. That is, we have asked the students to operate on, as an individual, this concept that was originally introduced as a collection of individuals.

When we talk about visualizations of distributions, we might want students to visualize the individuals in the original distribution being selected into the sample. Then, they need to see the sample mean becoming an individual in the sampling distribution. Let us look at a diagram that might depict these ideas.

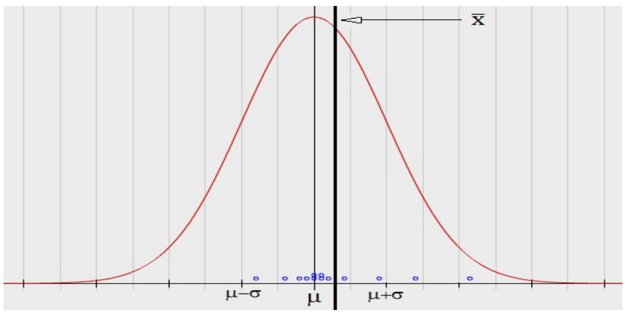


Diagram 3a: A normal distribution with a sample of 13 plotted and the mean of that sample identified.

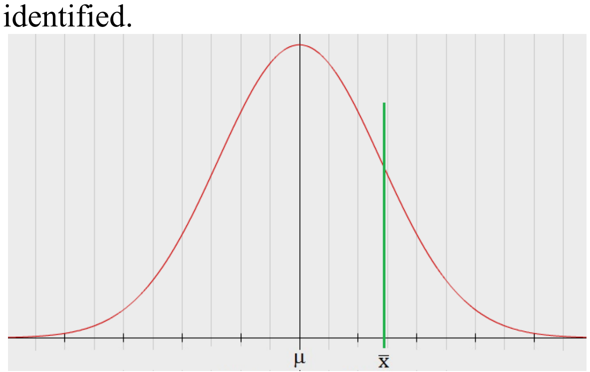


Diagram 3b: The distribution of all sample means (of size 13) from a normal distribution with the sample mean of the 13 points from Diagram 3a shown.

Summary and Future Directions.

There are two principle contributions of this study, both of which suggest directions for future research. The first contribution is that operating in the context of semiotics and by drawing upon Sfard's (1991) process-object distinction as well as the additional constructs of interiorization, condensation and reification we have outlined a hypothetical learning trajectory for the symbolic sense of mean, standard deviation and standard error of a sampling distribution. We have chosen these three topics because previous research has identified them as particularly difficult for students (Kim, Fukawa-Connelly, Cook, 2012; Mayen, Diaz, Batanero, 2009) and have important ties to more advanced statistical topics. This hypothetical learning trajectory is primarily based upon our experience as faculty; this theory needs empirical testing and revision. We will detail below some studies that we will carry out.

The second principle contribution of this study is to adapt Arcavi's (1994) description of a symbol sense to the realm of statistics. In doing so, we had to adapt Arcavi's notions to the particular domain of an introductory statistics class and we proposed some new ways of thinking that extend from Arcavi's into more specifically statistical ways of thinking. We have described the type of actions that students with a symbol sense would take and have drawn on our hypothetical learning trajectory to identify when, in the learning trajectory, students might be able to enact those behaviors in a meaningful way.

The most important aspect of testing our hypothetical learning trajectory relates to the importance of symbol sense in students' understanding. The reform statistics curriculum takes as an underlying assumption that students can develop a robust understanding of the concepts of statistics without much focus on the symbolic aspect, and that students can completely avoid the process-phase of symbolic understanding. In short, students, essentially, never calculate a standard deviation except using technology. As a result, if they are to develop an object-understanding of a standard deviation, it can only happen if they have simply skipped the process-understanding. We hypothesize that because undergraduate students have significant symbolic experience from their K-12 schooling that they are able to do this. That is, students' symbolic sense from their K-12 mathematics classes are able to compensate for the lack of computation we ask of our students. We hope to tease out the ramifications of this position by exploring students' symbolic understanding in mathematics and if they are able to and do transfer that understanding to the statistical context (we have some, minimal, evidence that they do, c.f., Kim, Fukawa-Connelly, & Cook, 2012).

The second area for exploration is to check the hypothetical learning trajectory itself. We will carry out studies to better understand how students' statistical sense develops and seek to confirm or refute our proposed progression. Once we have begun that process, we will also attempt to understand what type of statistical symbol sense is necessary for students to be able to develop robust linkages with the verbal and visual representations rather than merely memorizing the symbols. That is, what types of statistical understandings are necessary for students to be able to explain why a particular symbolic representation is appropriate and how it is linked to their (the students' personal) verbal and visual representations.

Finally, we mean to start a discussion about when it is appropriate for students to be focused on symbolic understanding? That is, when is it important for students to start developing a symbolic understanding? We do not believe that the statistics for non-math majors class is necessarily that place, but is an introductory class for majors the right place? We do assert that majors should be exposed to and expected to develop a robust proficiency with statistical symbols early in their program. We end with the question, "when should our majors develop their statistical symbol sense?"

References

Arcavi, A. (1994). Symbol sense: informal sense-making in formal mathematics, For the Learning of Mathematics 14(3), 24-35.

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.

Eco, U. (1976). *A Theory of Semiotics*. Bloomington: Indiana University Press.

Font, V., Godino, J., & D'Amore, B. (2007). An onto-semiotic approach to representation in mathematics education. *For the learning of mathematics, 27*(2), 2-7,14.

Garfield, J. & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. New York, NY: Springer.

Hewitt, D. (1999). Arbitrary and necessary: Part 1 a way of viewing the mathematics curriculum. *For the learning of mathematics, 19*(3), 2-9.

Hewitt, D. (2001a). Arbitrary and necessary: Part 2 assisting memory. *For the learning of mathematics, 21*(1), 44-51.

Hewitt, D. (2001b). Arbitrary and necessary: Part 3 educating awareness. *For the learning of mathematics, 21*(2), 47-59.

Kaplan, J., Fisher, D., & Rogness, N., (2009). Lexical ambiguity in statistics: How students use and define the words: association, average, confidence, random and spread. *Journal of statistics education*, *18*(1). http://www.amstat.org/publications/jse/v18n2/kaplan.pdf

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.

Lunsford, M., & Poppin, P. (2011). From research to practice: Basic mathematical skills and success in introductory statistics. *Journal of Statistics Education, 19*(1). http://www.amstat.org/publications/jse/v19n1/lunsford.pdf

Mayen, S., Diaz, C., Batanero, C. (2009). Students' semiotic conflicts in the concept of median. *Statistics education research journal*, *8*(2), 74-93.

Noth, W. (1995). *Handbook of semiotics (Advances in semiotics)*. Bloomington, IN: Indiana University Press.

Peters, S. Robust understanding of statistical variation. *Statistics education research journal*, *10*(1), 52-88.

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.

Watson, J., (2009). The influence of variation and expectation on the developing awareness of distribution. *Statistics education research journal*, *8*(1), 32-61.

Watier, N., Lamontagne, C., & Chartier, S., (2011). What does the mean mean? *Journal of statistics education*, *19*(2). http://www.amstat.org/publications/jse/v19n2/watier.pdf

Zieffler, A., & Garfield, J. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics education research journal*, *8*(1), 7-31.

GUIDED REINVENTION IN RING THEORY: STUDENTS FORMALIZE INTUITIVE NOTIONS OF EQUATION SOLVING

John Paul Cook
University of Oklahoma

The literature is replete with evidence of student difficulty in abstract algebra. In response, innovative approaches for teaching group theory have been developed, yet no corresponding methods exist for ring theory. In an effort to simultaneously fill this void and build upon Larsen's (2009) guided reinvention efforts in group theory, I conducted a study to investigate how students might be able to reinvent fundamental notions from introductory ring theory. Rooted in the theory of Realistic Mathematics Education, this paper reports on a teaching experiment conducted in nine sessions (up to 120 minutes each) with two students, neither of whom had prior exposure to abstract algebra. Using the construct of an emergent model, I show how these students formalized their intuitive understandings of linear equation solving and used them to reinvent the definitions of ring, integral domain, and field. In particular, the milestones of the reinvention process are identified and explicated.

Key words: Abstract algebra, ring, guided reinvention, Realistic Mathematics Education

Introduction and Research Questions

Rings are central structures in mathematics and enjoy an important place in the undergraduate mathematics curriculum. For typical mathematics majors, ring theory not only serves as the culmination of their mathematics careers but also lays a foundation for future study of advanced mathematics. Indeed, a solid understanding of the fundamental notions of ring theory is crucial for those students who wish to continue their study of mathematics in graduate school. Additionally, future mathematics teachers have much to gain from ring theory as it provides an underlying context for the techniques and axioms used in school algebra.

Despite the importance of rings, both in mathematics in general and in the undergraduate curriculum, there is reasonable evidence which suggests that students struggle mightily with the subject. While there are no studies which examines student difficulty with specific ring theoretic concepts, the literature is replete with evidence of students failing to understand even the most basic concepts in group theory (Dubinsky, Dautermann, Leron, & Zazkis, 1994; Hazzan & Leron, 1996). As rings are similar to, yet arguably more complex than, groups, it is quite reasonable to suggest that students experience rings with a comparable amount of difficulty. Compounding this issue is the fact that research addressing student learning of rings is almost nonexistent. In fact, only one study can be found in the literature (Simpson & Stehlikova, 2006). Thus, there is a considerable disparity between the significance of rings and the amount of information available to address student troubles with them.

In response to their own assertion that "the teaching of abstract algebra is a disaster, and this remains true almost independently of the quality of the lectures" (p. 227), Leron and Dubinsky (1995) suggested developing discovery-based methods for teaching the subject as an alternative to the traditional lecture. They proposed an investigative approach to instruction using the computer programming language ISETL. Additionally, using the theory of Realistic Mathematics Education (Freudenthal, 1991), Larsen (2004, 2009) developed an instructional theory which supports the guided reinvention of the concepts of group and group isomorphism. These efforts have since been expanded to create a complete reinvention-based curriculum for group theory (Larsen, Johnson, Rutherford, & Bartlo, 2009;

Larsen, Johnson, & Scholl, 2011). However, there are still no corresponding innovative instructional methods in the literature for ring theory. This study aims to begin filling this void by building upon Larsen's reinvention efforts in the arena of ring theory. In particular, this paper reports on a teaching experiment with two students designed to investigate how they might be able to reinvent the definitions of ring, integral domain, and field. The teaching experiment and its corresponding results are part of my larger dissertation project, wherein the ultimate goal is to constitute an instructional theory supporting the guided reinvention of these definitions. My research questions are as follows:

- How might students reinvent the definitions of ring, integral domain, and field?
- What models and activities are involved in developing these concepts when the students start with their own reasoning and intuition?
- What models and activities enable students to see the need for, define, and differentiate between additional ring structures like integral domain and field?

Literature

Of particular interest to this project is Larsen's (2004) dissertation wherein he produced an instructional theory supporting the guided reinvention of the definition of group. Using a developmental research design (Gravemeijer, 1998), he conducted three iterations of the constructivist teaching experiment (Cobb, 2000; Steffe, 1991) with two students apiece as a means of testing and revising his instructional theory. His instructional tasks centered on student manipulation of the symmetries of a triangle (and eventually other polygons). His students gradually formalized their intuitive notions with these symmetries and used them to write a precise mathematical definition of group. Larsen's dissertation and subsequent work established that the methods of guided reinvention are able to be used quite effectively in abstract algebra. Seeking a similar goal in the arena of ring theory, I adopted a similar theoretical perspective and research design (detailed in the methods section).

Like Larsen's work, nearly all of the literature concerning abstract algebra involves only group theory. Fortunately, the group theory literature does prove somewhat helpful, as rings and groups are structurally similar (in fact, a ring is an additive group with an additional multiplicative structure). Several features and learning mechanisms for groups which have been explored in the literature have direct analogs in ring theory. Those involving the definition of ring or the ring structure include, for example, binary operation (Brown, DeVries, Dubinsky, & Thomas, 1997; Iannone & Nardi, 2002), student proficiency (or lack thereof) with the group axioms (Dubinsky et al., 1994), confusion of the associative and commutative properties (Findell, 2000; Larsen, 2010), and the use of operation tables (Findell, 2000). Despite any possible application of this knowledge to student learning of rings, however, even introductory ring theory possesses several key, nontrivial features for which there is no analog in group theory: zero divisors, an additional binary operation, and the distributive property (to name a few). Information regarding these concepts can only be obtained by research which directly examines student learning of rings.

The lone article found in the literature which directly addresses student learning of ring theory is Simpson and Stehlikova's (2006) case study of how one student came to understand the commutative ring Z_{99}. This case study was used to draw conclusions regarding how students "apprehend" mathematical structure, defined as the shift of attention from the objects and the operations to the interrelationships between the objects as a result of the operations. The study examined the process by which a female student, Molly, apprehended a ring isomorphic to Z_{99} for her undergraduate thesis over a period of three years. It is worth noting that Molly had previously taken courses in abstract algebra, and consequently the researchers used a ring isomorphic to Z_{99} so that Molly would not immediately connect it with her prior, formal knowledge. Molly's primary self-guided method of apprehending this structure involved solving basic linear and quadratic equations. In addition to elucidating the need for the traditional ring axioms, this activity illuminated several key aspects of the ring structure: the existence of inverse operations, zero divisors, and units. These features arose as she attended to "the sense of interrelationships between the objects caused by the operations." Despite her three years of work with this structure, Molly never identified it as Z_{99}, nor did she exhibit any signs of accessing any of her formal knowledge from abstract algebra. Thus, it is reasonable to conclude that she used equation solving as method of discovering this ring structure (instead of a method of affirming what she already knew to be true about the structure).

In relation to my project, this case study suggests that equation solving can be used quite effectively by students in order to explore and apprehend an unfamiliar algebraic structure. In fact, this conclusion agreed with Kleiner's (1999) commentary on the historical role of equation solving in the rise of the axiomatic definition of a field: "In the solving of the linear equation $ax+b=0$, the four algebraic operations come into play and hence implicitly so does the notion of a field" (p. 677).

Theoretical Perspective

I adopted Realistic Mathematics Education (RME) as a theoretical perspective which guided both the instructional design and the data analysis. Two RME heuristics, in particular, were of critical significance to this study. First, the principle of *guided reinvention* (Freudenthal, 1991) served as the overarching guide for the study. The reinvention principle seeks "to allow learners to come to regard the knowledge they acquire as their own private knowledge, knowledge for which they themselves are responsible" (Gravemeijer & Doorman, 1999, p. 116). Secondly, the notion of an *emergent model* (Gravemeijer, 1998) was integral to the design of the instructional tasks and was used to identify milestones of the reinvention process. The purpose of an emergent model is to mediate a shift between informal mathematical activity to a new, more formal mathematical reality. The model is said to emerge as a *model of* the student's informal mathematical activity while gradually developing into a *model for* more formal mathematics. This process is known as the model-of to model-for transition. Gravemeijer (1999) delineated this transition into four phases of mathematical activity:

1. The *situational* phase involves working to achieve mathematical goals in an experientially real context.
2. The *referential* phase includes models-of that refer to previous activity in the original task setting.
3. The *general* phase is characterized by models-for that support interpretations independent of the original task setting.
4. The *formal* phase entails student activity that reflects the emergence of a new mathematical reality.

I viewed these phases as a continuous progression wherein activity within one phase would gradually progress toward the next. Because of the tendency for informal procedures to "anticipate" the emergence of more formal mathematical reasoning (Streefland, 1991), I argue that the progressive formalization within each phase *anticipates* the next. This expansion of Gravemeijer's four phases, then, can be expanded (if needed) to accommodate more detail by inserting three sub-phases. Namely, I introduce and define the following:

- The *situational anticipating referential* phase involves activity still firmly rooted in the original situational setting that lays the groundwork for future referential activity.
- The *referential anticipating general* phase is characterized by models-of that provide an overview of previous work in preparation for abstract or general activity.

- The *general anticipating formal* phase includes models-for which promote more efficient or concise use of the mathematics at hand in preparation for formal use.

I used these seven phases as a lens through which I present the results of the teaching experiment and identify the significant milestones of the reinvention process. This, in turn, provided a means by which I can begin to answer my research questions. Furthermore, it informs the creation of the emerging instructional theory being developed to support the reinvention of ring, integral domain, and field.

For the purposes of this project, I am viewing equation solving as an emergent model. Specifically, I anticipated that solving equations would initially serve as a model-of the students' informal activity with the ring structure, and that this would gradually transform into a model-for defining the desired ring structures.

Methods

I employed a *developmental research* design (Gravemeijer, 1998), which was compatible with and followed from my theoretical perspective because the primary goal is "the constitution of a domain specific instructional theory for realistic mathematics education" (p. 278). Following Gravemeijer's (1995) suggestion that the teaching experiment methodology is useful for such a purpose, I adopted the guidelines of the *constructivist teaching experiment* (Cobb, 2000; Steffe, 1991; Steffe & Thompson, 2000). In the constructivist teaching experiment, the researcher serves as the teacher and interacts with the students individually or in small groups (Cobb, 2000). I worked together with two students in the teaching experiment, which consisted of 9 sessions of up to 2 hours each.

Participants

The participant pool included students who had recently completed a course in discrete mathematics at a large comprehensive research university. Potential participants were recruited on a volunteer basis. At this university, the discrete mathematics course doubled as an introduction to advanced mathematics course and, aside from the course content, focuses on proof construction. To ensure the validity of the reinvention process, I wanted the participants to have had no direct prior exposure to abstract algebra, including group theory. I did require, though, that they had a working knowledge of modular arithmetic, polynomials, and matrices. Their familiarity with these concepts was assessed in a pre-survey administered after they had volunteered for participation. In addition to meeting the stated requirements, they were chosen on the basis of perceived compatibility with me and each other. I wanted two above average students and, ideally, one male and one female. The following table includes information on the two selected participants, Jack and Carey (pseudonyms):

Participants	Age	Major(s)	Discrete Math Grade
Jack	21	Mathematics	B
Carey	19	Mathematics & Physics	B

Instructional Tasks

Due to its potential for explaining the ring structure (Kleiner, 1999; Simpson & Stehlikova, 2006), solving linear equations became the focal point of the instructional tasks. Specifically, activities were designed that would culminate in solutions to additive and multiplicative "cancellation" equations $x+a=a+b$ and $ax=ab$ (a nonzero), respectively. Throughout the rest of the paper, I suppress the "a nonzero" qualifier so as not to detract focus from the two equations. I used the equation $x+a=a+b$ instead of the traditional $x+a=b+a$ to eliminate any ambiguity regarding the necessity of the additive commutativity axiom, which can be derived from the other ring axioms in a ring with identity (Dummit & Foote, 2004). These equations were chosen for their potential to both justify the ring axioms and enable students to differentiate between ring, integral domain, and field. For example, $x+a=a+b$ can be solved on an algebraic structure if and only if its additive structure forms an abelian group. The different methods of solving $ax=ab$ make use of all of the multiplicative ring axioms aside from commutativity (including multiplicative inverses). Additionally, $ax=ab$ serves to distinguish rings from integral domains, and integral domains from fields: it has a unique solution ($x=b$) if and only if the structure is an integral domain. In fields, this may be shown using multiplicative inverses or the zero-product property. On the other hand, in integral domains that are not fields it may only be proved by the zero-product property.

The structures upon which the specific linear equations and the cancellation equations would be solved were selected to incorporate examples of rings (that are not integral domains), integral domains (that are not fields), and fields so that each set of examples would be distinct in a meaningful way from the others. The structures I chose for the instructional tasks are the integers modulo 12, integers modulo 5, integers, polynomials in one indeterminate over the integers, and 2x2 matrices over the integers (throughout this paper, assume that these structures are accompanied by their usual operations):

Structure	Z_{12}	Z_5	Z	$Z[x]$	$M_2(Z)$
Rationale	finite, includes zero divisors	example of a finite field	prototypical ring structure; integral domain that is not a field	prototypical ring structure; integral domain that is not a field	prototypical noncommutative ring, includes zero divisors

Notice that I only included one example of a field, and one that is likely to be unfamiliar to students, at that. I additionally neglected to include the more familiar examples of fields, such as the real or rational numbers, in this initial set of examples, opting instead for an example of a finite field with five elements. This was done purposefully, in line with Zazkis' (1999) recommendation that "working with non-conventional structures helps students in constructing richer and more abstract schemas, in which new knowledge will be assimilated" (p. 651). Additionally, I planned for the students to generate their own examples after solving equations on the structures I provided, anticipating that they would introduce the more conventional examples of fields themselves.

Results

Recall that I am using equation solving as an emergent model to support the guided reinvention of the definitions of ring, integral domain, and field. The results of the teaching experiment are revealed, then, through the lens of Gravemeijer's (1998) phases of the emergent model transition, along with the three intermediate phases I introduced in the theoretical perspective section. In order, these phases are: situational, situational anticipating referential, referential, referential anticipating general, general, general anticipating formal, and formal. It is worth noting that, due to the gradual process of formalization I attempted to foster during the sessions, many of the students' initial solutions or responses to instructional tasks were not necessarily complete (or even correct). Instead of only presenting the students' finished products, I have included snapshots from the various stages to provide the reader with some context and feel for the reinvention process.

1. Situational: solving specific linear equations on Z_{12}, Z_5 , Z , $Z[x]$, and $M_2(Z)$

In addition to being designed as the original task setting, I classified the solving of specific linear equations on these structures as situational because it involves the students working towards a mathematical goal in an experientially real context. The students were initially directed to solve specific equations on the given structures, both to familiarize themselves with the features of each structure and with the activity of equation solving. The following was presented as a solution to the equation $x+3=9$ on the integers modulo 12:

$$-3 = +9$$
$$x+3+9 = 9+9$$
$$x+12 = 6$$
$$x = 6$$

As this example was taken from one of their initial responses, the solution is not yet complete and ignores, for example, associativity of addition. At first, the left hand side of the equation on second line of the solution read as x+3-3. I inquired about what was meant by -3, since it was not yet a defined element of the set. The students responded by defining -3 to be +9, and wrote this above their solution. When I asked them how this might be done for all "negatives" in Z_{12}, Carey responded by constructing a "negative number line" (the "as seen on a clock" addendum refers to a previous instructional task designed to increase their familiarity with modular arithmetic by likening addition modulo 12 to clock arithmetic):

Negative Number Line
as seen on a Clock

-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
12	1	2	3	4	5	6	7	8	9	10	11	12

Thus, the solving of $x+3=9$ enabled students to recognize the need for additive inverses. Additionally, examining the solution above makes it clear that the students recognized on some level, if not formally, that 12 is the additive identity of Z_{12}. Next, the students were prompted to solve multiplicative equations on Z_{12}. In particular, I gave them the equations $5x=10$ and $4x=8$ with the idea that they would recognize that $x=2$ is a solution for both but is only unique for $5x=10$. A near-complete solution to $5x=10$ is on the left, and an attempt to solve $4x=8$ is on the right:

$5x=10 \Rightarrow$
$5x\cdot 5=10\cdot 5$ Assuming commmunitivity
$5\cdot 5\cdot x=2$
$1\cdot x=2$ By our operation table
$x=2$

$4x=8$
$4x-4=8-4$
$4(x-1)=4$

Interestingly, regarding $5x=10$, the students opted for multiplication on the right (which necessitates the use of commutativity of multiplication) instead of the simpler multiplication on the left. The solving of these equations brought several other ring axioms to the fore as well: multiplicative inverse, multiplicative identity, distributivity, and, even though it was not yet recognized at this point by the students, associativity of multiplication. Additionally, in their attempts to solve $4x=8$, the students recognized a conceptual difference between the elements 4 and 5. While Jack and Carey were struggling to find an algebraic way to solve $4x=8$, I asked them about their need for a different technique:

Jack: It only works for numbers that are not a factor of our base.
JP: Right. So what is it that doesn't work in this other case?
Jack: 4 times any number does not make it 1.

While a correct solution to $4x=8$ was not produced until later, it is significant in that the students noticed that not all multiplicative equations can be solved in the same fashion.

2. Situational anticipating referential: solving $x+a=a+b$ and $ax=ab$ on each of the given structures

This activity could be easily be classified as simply "situational," because these equations could have been the focus of the original task setting on their own. In other words, the students' ability to solve these equations could have been independent of solving specific equations beforehand. On the other hand, however, the specific equations were used as a paradigm upon which the students could reference to solve $x+a=a+b$ and $ax=ab$. Solving these general equations was also designed to promote the summarization of their previous activity, thus anticipating the need for referring to these results at a later stage. Consequently, I classified this activity in the intermediate stage of situational anticipating referential. The following excerpt was their near-completed solution to the equation on $Z[x]$ (they used capital letters to denote polynomials in x):

$$X + A = A + B$$

$$(X + A) + (-A) = (A + B) + (-A)$$

$$X + 0 = (A + B) + -A$$

$$X = (B + A) + -A$$

$$X = B + (A + -A)$$

$$X = B + 0$$

$$X = B$$

Setting aside the fact that additive associativity was omitted between lines 2 and 3 (though is used correctly throughout the rest of the solution), all of the additive ring axioms are in play here. By this point, they had written out a solution nearly identical to this one for the preceding three structures as well, prompting them to remark:

Jack: Adding [polynomials] is basically adding integers.
Carey: So you do the same thing that you did before.

A similar exchange occurred when they started to solve $x+a=a+b$ on $M_2(Z)$:

Jack: It is commutative, A + B = B + A.
Carey: You can do the same thing that you did for the addition, because you just add the complements.

JP: And they are [matrices] over the integers. So the components are commutative. So you guys are saying that it's exactly the same thing as the others?
Carey: Yeah.
Jack: Yes.

This dialogue suggests that, in addition to successfully motivating the need for all of the additive ring axioms, the equation solving model effectively highlighted the identical additive structure present in all rings. However, the multiplicative structure is a different story, and this was recognized at once by the students when they wrote up their solutions to $ax=ab$ for Z_{12} (left) and Z_5 (right). Recall that each unit is its own inverse in Z_{12}.

If $a \cdot a = 1 \Leftrightarrow a \in \{1,5,7,11\}$
$ax = ab$, given
$(ax)a = (ab)a$ multiplication
communitivity $(xa)\cdot a = (ba)\cdot a$
$x \cdot (a \cdot a) = b(a \cdot a)$ associativity
$x \cdot 1 = b \cdot 1$ assumed
$x = b$ by definition of $x1$
Q.E.D.
multiplicative ~~prop~~ id.

$ax = ab$ Given $a \neq 5$
MI $(ax)a^{-1} = (ab)a^{-1}$
Comm $(xa)a^{-1} = (ba)a^{-1}$
Asso $x(aa^{-1}) = b(aa^{-1})$
MI $x \cdot 1 = b \cdot 1$
MId $x = b$

The critical difference the students noticed here was that the solution in Z_5 was valid for all nonzero elements, whereas the solution in Z_{12} only held for a small subset. And while the above methods are similar, Jack and Carey noticed that they were not able to use this method in general when they were faced with solving $ax=ab$ over the integers,

Carey: Did we define division?
JP: What would happen if you did that?
Carey: Like x over a equals x times 1 over a.
Jack: The problem is what is this? 1 over a. It's not going to exist over the integers necessarily. That's not necessarily going to be in the integers unless it's 1.
JP: What else could it be?
Jack: Negative 1, I guess.

They opted to use distributivity with the zero-product property instead:

$ax = ab$ ← $ax - ab = ab - ab$
$ax - ab = 0$ Additive inverse
$a(x-b) = 0$ Distributive prop.
$\cancel{a}x - b = 0$ given that $a \neq 0$
$x = b$
$(x-b)+b = 0+b$
$x(-b+b) = b$
$x = b$
using:
$ab = 0$, $a=0$ or $b=0$
if $a \neq 0$
then $b = 0$

This solution, in addition to identifying the necessity of the distributive and zero-product property, also helped the students to mentally differentiate the integers (and then, eventually, polynomials) from the modular rings with which they had worked previously. Thus, the students' solving of $x+a=a+b$ and $ax=ab$ on each of the five structures:

1. reinforced the need for the axioms used to solve the specific equations,
2. enabled them to see that all of the examples had identical additive structures, and
3. enabled them to notice the differences in multiplicative structure.

3. Referential: summarizing the results from solving $x+a=a+b$ and $ax=ab$

After the equation solving activities were completed, I gave the students a task prompting the students to organize their solving of the equations $x+a=a+b$ and $ax=ab$. Specifically, they were asked to identify the different methods they used to solve the equations, and whether the given method could be solved always, sometimes, or never on each of the structures. I classified this task as referential because it was distinct from the original task setting yet referenced the previous activity in the original task setting. Additionally, at this point, the model is still a model-of their equation solving activity and had not yet transitioned into a model-for (which takes place in the general phase).

Once Jack and Carey had discussed the different methods for solving the equations on each of the examples, I had them organize their results in a chart by writing "A" for "always works", "S" for "sometimes works", and "N" for "never works" (across the top row: $x+a=a+b$; $ax=ab, a \neq 0$ using mult. inverses; $ax=ab, a \neq 0$ using distributivity and the zero-product property):

	$x+a=a+b$	$ax=ab, a\neq 0$ using mult. inverses	$ax=ab, a\neq 0$ using distributivity and zero-product property
$\mathbb{Z}_{12}$	A	S	S
$\mathbb{Z}_5$	A	A	A
$\mathbb{Z}$	A	S	A
polynomials over integers	A	S	A
2×2 matrices w/ integer entries	A	S	S

In addition to summarizing their previous work, the students were also required to build off of it. For example, the students had not yet considering whether $ax=ab$ could be solved on polynomials by using multiplicative inverses:

Jack: Polynomials over integers. [Multiplicative inverses] held, didn't it?
Carey: We didn't do it that way.
JP: What would happen if you tried to construct a multiplicative inverse for a polynomial? $1/x^2$. Is that a polynomial, based on how we defined it?
Carey: Basically, we are starting with n = 0, which we are.
Jack: So we can't do things with polynomials.

The discussion continued until they realized that 1 and -1 were the only polynomials which had multiplicative inverses, earning *Z[x]* a rating of "sometimes" in that column. Similar discussions were held for methods which had not yet been applied to other structures.

A number of interesting patterns emerged in the chart, both from my perspective and the students'. First, the students recognized that there is essentially only one way to solve the additive equation. Jack noticed this during the activity by referencing their previous work solving $x+a=a+b$, remarking, "I think that this method works in all of the cases." Second, notice that the sets with "identical ratings" do indeed have substantial features in common. The always-sometimes-sometimes rating appears for Z_{12} and $M_2(Z)$, which are the structures containing zero-divisors. The always-sometimes-always rating appears for Z and $Z[x]$, which are the integral domains that are not fields. Lastly, Z_5 and its always-always-always rating stands alone as the only field under consideration (at this time).

At this point, I encouraged the students to generate their own examples of structures upon which the given equations could be solved (in other words, sets endowed with addition and multiplication). Then I prompted them to fill out a similar chart for their new examples:

	$x+a=a+b$	$a\neq 0$ $ax=ab$ by mult. inverses	$ax=ab, a\neq 0$ distributivity and ZPP
$\mathbb{R}$	A	A	A
$\mathbb{C}$	A	A	A
$\mathbb{Q}$	A	A	A
$\{0\}$	A	N	N
$\mathbb{Z}_n$ $n\in\mathbb{N}$, n prime	A	A	A
$\mathbb{Z}_n$ $n\in\mathbb{N}$, n composite	A	S	S

As I had previously anticipated, the students' own examples were dominated by fields. In fact, four of the six student generated examples were fields (specifically, the real numbers, complex numbers, rational numbers, and integers modulo a prime). Notice also that they differentiated between Z_n for n prime and composite (this occurred before the chart activity as a result of generalizing their reasoning about Z_5 and Z_{12}). As expected, the fields and their always-always-always ratings agree with the ratings for Z_5 on the previous chart. The only example I had not anticipated was {0}, the trivial ring. Because this example is markedly different from the other examples and does not lend any insight into the ring structure, I intervened and removed it from further consideration (though, in the interests of using student-generated ideas as much as possible, I did re-introduce it after the definitions had been reinvented).

4. Referential anticipating general: the sorting activity

Now that Jack and Carey had organized the results of their equation solving, I gave them a sorting task to encourage them to sort the structures based on what they felt were common characteristics using their charts. I classified this as referential anticipating general because it involved referring to previous activity (the chart activity and, to a lesser extent, the actual equation solving activity). In this way, this task was not yet "independent of the original task setting", a characteristic of the general phase. Sorting based on common features, however, does anticipate the mathematical activity of abstraction, which certainly qualifies as general activity. Since the students performed the bulk of the mathematical activity for this task by filling out the charts, this activity proved to be quite simple. Jack commented "If we are not categorizing them by the first column, which is trivial, we are categorizing them by the second column and the third column," suggesting that the equation solving chart is now a model-of the identical additive structure for all rings (ratings in the $x+a=a+b$ column are all "always") as well as the differing multiplicative structure (differing ratings for the $ax=ab$ columns). This realization enabled the students to sort based on the ratings for $ax=ab$:

Group 1: $\mathbb{Z}_5$, $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{Z}_{n}$, n prime   Group2: $\mathbb{Z}$, polynomials   Group3: $\mathbb{Z}_{12}$, matrices, $\mathbb{Z}_n$, n: composite

Thus, at this point, the students have used equation solving as a means to sort these structures. Jack's comment above emphasizes that the primary criteria for sorting included how $ax=ab$ can be solved on each of the structures. The underlying ring-theoretic concepts which govern how this equation be solved, of course, are the existence (or lack) of zero-divisors and multiplicative inverses. Whether the students were formally aware of these features at the time of the sorting is unclear. It is clear, however, that the ratings for different methods of solving $ax=ab$ on each structure served as a model-of these ideas for the students.

5. *General: defining by abstracting common features*

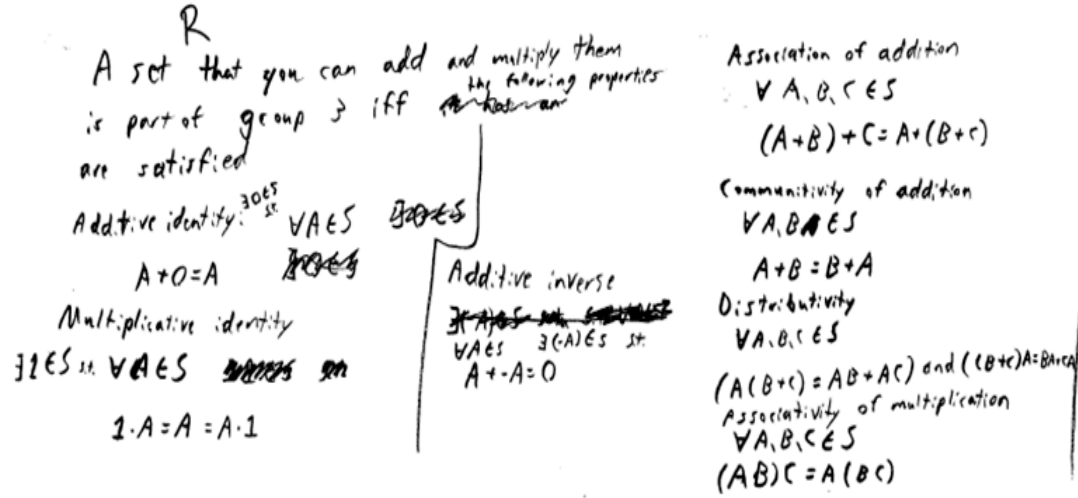
At this point, I asked the students to define a list of criteria for inclusion in each of the three sets. This required them to identify the common characteristics of each collection. This activity was classified as general because, finally, the equation solving model had emerged as a model-for the formal activity of defining the different ring structures, independent of the original situational task setting. Again, I asked them to display their results in a chart by listing the rules they had used to solve the equations and determining if the given rule holds for each group. The chart is reproduced here:

	Group 1 R, C, Q, Z_5, Z_n for n prime	Group 2 $Z, Z[x]$	Group 3 Z_n for n composite, $Z_{12}, M_2(Z)$
Additive identity	X	X	X
Multiplicative identity	X	X	X
Associativity of addition	X	X	X
Commutativity of addition	X	X	X
Distributivity	X	X	X
Zero-product property	X	X	
Associativity of multiplication	X	X	X
Commutativity of multiplication	X	X	
Multiplicative inverse	X		
Additive inverse	X	X	X

This served as a springboard to begin the process of defining. I followed Larsen's (2004) guidelines for supporting a cyclic process of presenting and revising a definition:

1. The students prepared a definition.
2. I read and interpreted the definition, calling attention to particular choices made by the students.
3. The students revised their definitions as necessary and restarted the process.

I suggested that they start with group 3, the idea being that starting with what I knew to be the most general structure would provide them with the possibility of defining subsequent structures in terms of this one. Here is one of Jack and Carey's initial attempts to write out the criteria for a structure to be included in group 3:



This, of course, is a preliminary definition of a ring with identity. At this stage of the defining process, the students still need to address the existence of the binary operations and issues with quantifiers, among other things. After this definition was completed, the students wrote a similarly rough definition for group 2. It was at this point that I gave them the names for the structures in each of the three groups so that they could finalize their formal definitions. In writing their initial definition of an integral domain, Jack and Carey did not immediately see the potential for defining integral domain in terms of a ring with identity.

6. General anticipating formal: writing "nested" definitions

When the students were repeating their definition of integral domain to incorporate revisions, the students notified me of what a mundane process rewriting the same axioms would be. I used this as an opportunity to engage them in a conversation about how they could shorten the process:

> JP: So as you guys have correctly noted writing all of these out is a huge [inconvenience], so if we wanted to write out, say the next one, knowing that we have this definition down now, what's a way that we could shorten, shorten the next one.
> Jack: We just say, if it's A ring, and has the following properties.
> JP: Okay. So, how would you do that? …
> Jack: Uh, oh if you wrote the main rings then the difference between a ring with identity is that [an integral domain] has a few more properties.
> JP: Okay. Yeah.

As a result, Jack and Carey wrote a definition of an integral domain in terms of a ring with identity. Shown is their finalized version of this definition:

A set R with $+:R\times R\to R$ and $\cdot : R\times R\to R$
is an integral domain iff it is a ring (with iden
and has the following properties:

Zero product property: Commutativity of multiplication

$\forall A,B \in R$ $\forall A,B\in R$ $AB=BA$

$AB=0 \Rightarrow a=0 \vee b=0$

They used the same technique to define a field in terms of an integral domain:

A set R with $+:R\times R\to R$ and $\cdot : R\times R\to R$ is a field iff it is an integral domain & satisfies the following ~~propo~~ property:

Multiplicative inverse:

$\forall A\in R,\ \exists A'\in R$ s.t.

s.t. $A\neq 0$,

~~does~~

$A(A')=1$

I categorized this activity as general anticipating formal because it still involves the defining of mathematical structure (which I previously argued is general), while the "nesting" of these definitions served as a tool for classifying other ring structures and emphasizes the interrelationships between the three definitions. Thus, nesting the definitions prepared the definitions for their use in a more formal setting.

7. *Formal: using the reinvented definitions to classify other examples of rings*

Upon the reinvention of the definitions of ring with identity, integral domain, and field, I turned the students' attention to tasks in which they would use the definitions to classify other examples of rings. These tasks qualify as formal as they reflect the emergence of a new mathematical reality. Specifically, one of the tasks asked the students to classify $Z_3 \times Z_3$ (with the usual component-wise operations modulo 3). I anticipated that they would initially conjecture that it is a field (since Z_3 is a field), and that they would find this to not be the case. Indeed, after verifying that all of the axioms for a ring with identity (plus multiplicative commutativity) held, they turned their attention to the zero product property and multiplicative inverses. When examining the zero-product property, I named and defined the term "zero divisor", a concept with which they were familiar at this point due to their experience with Z_{12} and $M_2(Z)$. I then asked them if there were any zero divisors present in $Z_3 \times Z_3$. In a similar fashion, I asked Jack and Carey which elements had multiplicative inverses (and named these, accordingly, as units). These were the results (zero-divisors are on the left and units are on the right; note that $Z\times Z$ should be $Z_3\times Z_3$):

$Z\times Z$

$\langle 0,a\rangle\cdot\langle b,0\rangle=\langle 0,0\rangle$

$a,b\neq 0$

So zero-product prop. doesn't hold.

$\langle a,b\rangle\cdot\langle c,d\rangle=\langle 1,1\rangle$

$ac=1$

$bd=1$

everything $\langle a,b\rangle$ where $a \& b \neq 0$

Thus, they concluded that $Z_3 \times Z_3$ is a ring with identity that also has a commutative multiplication (I use this opportunity to introduce the notion of a *commutative ring*). Then I asked the students to classify the infinite ring $Z \times Z$. A conversation ensued regarding the zero product property:

> Jack: Z cross Z would be…
> Carey: Don't we have a similar problem?
> Jack: It would still be a ring.
> Carey: Yeah.
> Jack: It would have the exact same problem with zero-product property 'cause there's going to be…you can just take pairs of zeros out of it.

They concluded that, since the zero-product property did not hold, that $Z \times Z$ could not be an integral domain or a field. This excerpt, in addition to displaying the students' activity in a new mathematical reality, demonstrates that the students having a functional, working knowledge of the definitions they reinvented.

Conclusions

In addition to providing information about how students come to understand fundamental concepts in ring theory, this paper supports two primary conclusions which contribute to the knowledge of the field. First, I introduced an expansion of Gravemeijer's (1999) phases of the emergent model transition based on the idea of *anticipation* and progressive formalization. Using the results of a teaching experiment designed to investigate how students might able to reinvent the definitions of ring, integral domain, and field, I presented evidence that demonstrates how such an expansion can be useful when explaining and interpreting the emergence of a model.

Second, the results of the teaching experiment demonstrate how students might be able to capitalize on their informal knowledge of solving equations in order to reinvent the fundamental ring structures. Additionally, the adaptation of Gravemeijer's model to seven phases highlighted the significant milestones of the reinvention process, laying the groundwork for a domain-specific instructional theory to support the reinvention of these definitions:

1. Solving specific linear equations on a variety of ring structures
2. Solving the equations $x+a=a+b$ and $ax=ab$ on each of the structures
3. Summarizing the different methods used to solve $x+a=a+b$ and $ax=ab$
4. Sorting the structures based on similar methods used to solve $x+a=a+b$ and $ax=ab$
5. Defining by abstracting the common features of each set of sorted structures
6. Writing "nested" definitions, i.e. writing specific definitions in terms of general ones
7. Using the definitions for more formal activity (such as classifying other rings)

This emerging instructional theory framework will be tested, refined, and elaborated through another iteration of the teaching experiment as a part of my dissertation research project.

References

Brown, A., DeVries, D., Dubinsky, E., & Thomas, K. (1997). Learning binary operations, groups, and subgroups. *Journal of Mathematical Behavior, 16*(3), 187-239.

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. Kelly & R. Lesh (eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 307-334). Mahwah, NJ: Erlbaum.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics, 27*(3), 267-305.

Dummit, D. & Foote, R. (2004). *Abstract Algebra*. Hoboken, NJ: Wiley.
Findell, B. (2000). *Learning and understanding in abstract algebra*. Unpublished Doctoral Dissertation, The University of New Hampshire.
Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht, The Netherlands: Kluwer.
Gravemeijer, K. (1995). *Developing realistic mathematics instruction*. Utrecht, The Netherlands: Freudenthal Institute.
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. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1*(2). 155-177.
Gravemeijer, K., & Doorman, M. (1999). Content problems in realistic mathematics education: a calculus course as an example. *Educational Studies in Mathematics*, 39, 111-129.
Hazzan, O., & Leron, U. (1996). Students' use and misuse of mathematical theorems: The case of Lagrange's theorem. *For the Learning of Mathematics, 16*(1), 23-26.
Iannone, P. & Nardi, E. (2002). A group as a 'Special Set'? Implications of ignoring the role of the binary operation in the definition of a group. *Proceedings of the 26th Annual Conference of the International Group for Psychology in Mathematics Education*, (3), 121-128, Norwich, United Kingdom.
Kleiner, I. (1999). Field theory: From equations to axiomatization. *American Mathematical Monthly, 106*(7). 677-684.
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. (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. (2010). Struggling to disentangle the associative and commutative properties. *For the Learning of Mathematics, 30*(1), 37-42.
Larsen, S., Johnson, E., Rutherford, F., & Bartlo, J. (2009). A local instructional theory for the guided reinvention of the quotient group concept. *Proceedings of the 12th Annual Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC.
Larsen, S., Johnson, E., & Scholl, T. (2011). Putting research to work: Web-based instructor materials for an inquiry oriented abstract algebra curriculum. *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education*. Portland, OR.
Leron, U. & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly, 102*(3), 227-242.
Simpson, S. & Stehlikova, N. (2006). Apprehending mathematical structure: a case study of coming to understand a commutative ring. *Educational Studies in Mathematics*, 61, 347-371.
Steffe, L. (1991). The constructivist teaching experiment: Illustrations and implications. In E. von Glasersfeld (Ed.), *Radical Constructivism in Mathematics Education* (pp. 177-194). The Netherlands: Kluwer.
Steffe, L. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A. Kelly & R. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 267-306). Mahwah, NJ: Erlbaum.
Streefland, L. (1991). *Fractions in realistic mathematics education: A paradigm of developmental research.* Dordrecht, The Netherlands: Kluwer.

FUTURE TEACHERS' INTENTIONS FOR GENDER EQUITY: HOW ARE THESE CARRIED FORWARD INTO THEIR CLASSROOM PRACTICE?

Jacqueline M. Dewar and Rozy Vig
Loyola Marymount University

Mathematics teachers at all levels are called to promote gender equity in their classrooms. During a college course on mathematics and gender, future K-12 teachers indicated their intentions to foster gender equity in their own classrooms. To investigate whether, and how, this resolve for equity persisted and influenced their own classroom practice, we present case study data of four former students from this course. Using a grounded approach (Glaser, 1992) to analyze classroom observations and semi-structured interviews, we report how closely the former students' current descriptions of an equitable classroom align with their classroom practice, and with NCTM's call for equity. We find that these teachers' self-assessment of their success in achieving equitable classrooms appears to be accurate. We also highlight the learning experiences they feel most contributed to their views and practice regarding equity and equitable teaching. The results suggest possible implications for mathematics teacher preparation programs.

Key words: [K-12 teacher preparation, gender equity, classroom practice, role model, case study]

Teachers are called to play a role in ensuring gender equity in mathematics instruction (Secada, Jacobs, Becker & Gilman, 2001; National Coalition for Equity in Education, 2003). The phenomenon known as stereotype-threat (Steele, Spencer, & Aronson, 2002) makes excelling in mathematics more challenging for female and minority students. Enlightening future teachers about the facts and fallacies that underlie the widely held idea that boys are simply better at math than girls (Halpern et al., 2007; Hyde, Lindberg, Linn, Ellis, & Williams, 2008) is one way to empower teachers to confront these stereotypes personally so that they can then help their students to do so in their own lives. Providing information about role models by fostering awareness that women have contributed to the development of mathematics is another important strategy for encouraging underrepresented groups in mathematics (Marx & Roman, 2002; Leonard, 2008; Bonetta, 2010). These ideas informed the design of a course on women and mathematics, which prompted the study of how future teachers' intentions to teach equitably are carried forward into their own classrooms.

A Brief Description of the Course

The course was originally developed and taught by the first author.[1] It was inspired by the publication of the book *Math Equals* (Perl, 1978). Like the book, the course examines the lives of nine women mathematicians from Hypatia in the fourth century to Emmy Noether in the twentieth century. It engages students in mathematical activities related to the work of those women, which allows for discussion of mathematical topics ranging from conic sections to functions to elementary group theory. To bring structure to the wide variety of mathematical topics, the course emphasizes three mathematical themes:

[1] Three rounds of funding from the Tensor-MAA Women and Mathematics grant program have supported the first author in team teaching this course in 2008, 2010, and 2012, each time with a different junior faculty member from her department. As part of the dissemination effort for this grant, a course webpage (http://myweb.lmu.edu/jdewar/wam) provides more information about the course.

1. Mathematics, at its heart, is a study of patterns and not numbers.
2. Inductive and deductive reasoning play distinct and vital roles in mathematics.
3. The use of multiple representations for a given concept can be valuable in learning and teaching mathematics.

The course attracts students in the elementary school credential program who want to concentrate in math but have had little or no calculus, as well as students majoring in mathematics who have taken a number of upper division math courses. To accommodate the range of students' math backgrounds, the instructor strives to approach these topics in novel and accessible ways. For example, when discussing conic sections in conjunction with Hypatia, all students can comprehend and appreciate the Dandelin sphere proof (of the equivalence of the "cutting-the-cone" and the "distance-focus" definitions of the ellipse) and rarely have students seen it prior to the course. The course also examines research on gender differences in mathematics participation and achievement as seen in the 1970s and today. In addition, it asks how the experience of contemporary women mathematicians, particularly those of color, intersects with the experiences of the nine women from history.

Research Questions

A previous study of the course (Dewar, 2008) focused on the change in students' views of mathematics. In an end-of-term portfolio the future teachers voluntarily pledged to encourage <u>all</u> their students equally in the study of math. This was significant because the reflection prompt they were addressing dealt with mathematics and <u>not with equity</u>. This prompted a new study of the course, which sought to determine:

- How does future teachers' resolve for an equitable classroom get carried forward into their classroom practice?
    - As teachers, what are their current views and actions toward gender equity?
    - What courses, learning experiences or pre-professional opportunities fostered these views or actions?
    - What factors support and hinder them in achieving gender equity in their classrooms?

Research on equity highlights two differing views with respect to classroom practice (Streitmatter, 1994). One is to provide equal opportunity (at the outset) with the assumption that differences in outcomes are a function of individual differences. The other is to aim for equal outcomes. Here the teacher provides additional resources to try to meet special needs or to compensate for disadvantages. Streitmatter (1994) acknowledges problems with both of these approaches. The first takes no account of different backgrounds, motivations, or beliefs resulting from past educational experiences or inequities, societal biases, or stereotype threat. The second, some feel, can result in reverse discrimination. NCTM comes down on the side of the second as seen in the Equity Principle in *Principles and Standards* (2000) and in the *Changing the Faces* series (2001). In this paper, we will show that the teachers in the study also come down on the side of the second view, aiming for equal outcomes in their teaching.

Methods

Subjects

We were able to identify and contact four former students who were teaching in the local area. They all agreed to participate in the study. All had taken the women and mathematics course in 2008 from the first author, a mathematics professor. All four are women, two are white, and two are Hispanic. One student majored in liberal studies and minored in mathematics. She was pursuing her elementary teaching credential, while seriously considering getting a secondary math credential. This was her second year of teaching first grade in a private Catholic elementary school. The other three students majored in mathematics and were each teaching at the high school level. One taught in a private Catholic school and the other two taught in public charter schools. The high school teachers were all in their third year of teaching. For the remainder of this paper, each teacher will be referred to as Instructor #1 (I#1), Instructor #2 (I#2), Instructor #3 (I#3), and Instructor #4 (I#4), with I#1 denoting the elementary school teacher.

Classroom Context

In the classrooms of I#1 and I#3, females comprised the majority of the students, while in the classrooms of I#3 and I#4, females were in the minority. Although the types of schools, grade levels, and the gender ratios varied across the four observed classrooms, all four schools had a student population consisting almost entirely of Hispanics and African-Americans. On the day of observation, all the classrooms were populated by either Hispanic or African-American students. The three secondary teachers were each observed teaching a geometry lesson. Table 1 contains a summary of the school and instructor characteristics, the gender data for each classroom, and the ethnicity data for each school.

Level of Instructor	**#1 first grade, #2, #3 and #4 secondary**			
	#1	**#2**	**#3**	**#4**
Private/Catholic School	**x**	**x**		
Public/Charter School			**x**	**x**
High School Geometry		**x**	**x**	**x**
1st Grade Math	**x**			
Teaching Experience	**2**	**3**	**3**	**3**
Classroom: Gender count	**12F:6M**	**13M:3F**	**15F:5M**	**15M:7F**
School: Ethnicity data[2]	**H: 99%**	**H: 28% AA: 62%**	**H: 99%**	**H: 96% AA: 3%**

Table 1 Characteristics of Instructor, School, and Classroom

Data

Each classroom was observed once. Notes were taken on classroom pedagogy, classroom discourse, and materials posted on the walls. The classroom pedagogy was observed for an overall sense of the learning environment as being more or less teacher-centered or student-centered. With regard to classroom discourse, to the extent possible, notation was made of the genders of the students who volunteered answers or were called on by the teacher. Based on these notes, a gender response ratio (the gender ratio of students who volunteered answers or were called on directly by name) was computed for each lesson and compared to the gender ratio of students in the classroom. We caution that care must be taken in interpreting such a comparison. As we will see in the actual data, the gender response ratio can appear to be fairly favorable to one gender while the actual gender count of responders differs from the actual gender count in the classroom by a single response. The materials posted on the classroom walls were observed for references to gender equity. For example, notes were taken of posters depicting male and female mathematicians and of the type of student work on display.

Semi-structured interviews, approximately 45 minutes in length, were conducted after each classroom observation. During the interview, the instructors were asked to describe an equitable classroom and discuss what factors supported and hindered their ability to achieve equity in their classrooms. All the instructors were asked: How important is it for you to

[2] This ethnicity data for the 2011-12 academic year was obtained from internet sources.

create an equitable classroom, very important, important, somewhat important, not at all important, or you never think about it? This question was followed by an open-ended question to inquire about the instructors' views of gender equity: What comes to mind when you think of an equitable classroom? The instructors were then given a series of follow-up questions: Would you say equity means *equal* or *something different*? How are you attending to issues of gender equity in your classroom? Do you feel you have been successful in creating a gender equitable classroom? Why or why not? These conversations led naturally to the questions: What factors support your ability to achieve equity in your classroom? What factors hinder your ability to do so? What educational and pre-professional experiences have shaped your views on equity?

Data Analysis

All four interviews were audiotaped and transcribed for analysis using grounded theory with open coding (Glaser, 1992). Both authors independently read one transcript and identified similar themes. The first author then developed a coding scheme and coded all four transcripts to arrive at the constructs of student voice (SV) and role models (RM). Below is a description of the analytic process used to arrive at the SV and RM constructs.

Following a thorough reading of all the transcripts, participation, confidence, and engagement emerged as the most frequently mentioned themes. These themes were developed into a coding scheme. Codes P, C, and E were used to denote references to student behavior related to participation, confidence, and engagement, respectively, while the codes IP, IC, and IE referred to strategies the instructors used or could use to promote these behaviors. Statements about behavior in the service of verbal participation and strategies that promote verbal participation were coded P and IP, respectively. Statements referring to gender differences in attitudes and beliefs about oneself as a doer of mathematics were coded C. References to strategies for addressing such differences were coded IC. Statements about interests in and motivation to work were coded E, while any mention of strategies for promoting interest and motivation were coded IE. Below are examples for each of the codes:

P - "The girls raise their hands less."
IP – "I consciously make an effort to call on a girl, call on a boy."
C – "The girls weren't positive, they always questioned their answers."
IC – "I make a very strong point that girls can do math in my class."
E – "You could tell he was into his work."
IE – "Teaching students to self-assess gets them more engaged."

Codes were chunked when appropriate. For example, if an instructor made two back-to-back statements about promoting verbal participation with little to no difference in content, both statements were chunked together and received a single IP code. If however, the second statement introduced a new strategy or new idea with respect to participation, the two statements each received an IP code. Finally, all of the above codes were gathered into a single construct called "Student Voice" (SV) which was used to frame the results.

The importance of Role Models (RM) emerged as another frequently mentioned theme and was considered a second construct of interest. The instructors referenced RM across a number of contexts: as a recognition of how few RM the instructors themselves had encountered, as a concern that their students should have RM, as a thought that they themselves might serve as a RM for their students, and/or as a strategy to promote equity.

The codes mentioned above are by no means exhaustive with respect to the data. Both I#2 and I#4 made reference to giving equally challenging problems to males and females as a strategy to promote equity. This was one theme that emerged from the data but did not fit into either the SV or the RM construct. Because the themes specific to SV (i.e., participation, confidence, and engagement) and RM were the most commonly mentioned by all four instructors, we focus our attention on those for the remainder of the paper.

The constructs of SV and RM that emerged from the interview data were compared with what was observed in their classroom practice. In addition, physical manifestations of SV were noted in the classroom (student work displayed, motivational sayings or posters, and posters showing applications of mathematics), as were physical manifestations of RM (posters of female or minority mathematicians).

Results

In what follows, we will report on classroom observation data and interview data for each instructor. The classroom data will highlight the pedagogy and discourse used in the classroom as well as the materials found on the walls. The interview data will bring to light influences that prompted each instructor to strive for equity, factors that supported and hindered their success, and self-assessments of their achievements.

Instructor #1 Classroom Observation

I#1, the first grade teacher, employed a combination of age-appropriate teaching methods that involved whole group recitation, and individual work using hands-on manipulatives and worksheets. There was also some Socratic-type discourse with the whole class. She used a range of classroom management techniques to engage her students and keep them on task. For example, a song was used to transition between lessons. On some tasks, students (early finishers) were asked to help other students in their group. I#1 also incorporated reading into the math lesson. She used an informal assessment technique at the end of the lesson, called an "exit slip," which on this day consisted of a single subtraction problem. Females outnumbered males 2 to 1 in this class, and the gender response ratio 2.4 to 1 favored females. However, the gender response count (12 female voices to 5 male voices) only differed by one from the student gender count in the class (12 females to 6 males). The student work and progress charts displayed in the classroom represented a physical manifestation of SV. There were no physical manifestations of RM.

Instructor #1 Clinical Interview

I#1 indicated that equity was very important to her and stated, "An equitable classroom means each child has an equal opportunity to learn. To achieve that may require the instructor to parcel out time and attention unequally." She mentioned the woman and mathematics course as influencing her desire to achieve gender equity, in that it made her aware of mathematical role models for females, especially for herself. She cited the statistics on women's participation and achievement in mathematics discussed in the course as raising her awareness of gender inequities in mathematics. She had used teaching materials derived from the course once, but found that while the lesson on math as patterns was meaningful to her first grade students, they could not grasp the connection to a particular woman mathematician (Sonya Kovalevskaya). She also mentioned the math methods course as helping her develop equitable teaching methods. Regarding hindrances to achieving equity, I#1 noted that gender stereotypes held by some parents had led to complaints about her classroom policy that all students use wheeled bags rather than backpacks, but these were unrelated to mathematics instruction. She stated that she was generally satisfied with what she was achieving in her classroom in terms of equity.

Instructor #2 Classroom Observation

In I#2's secondary classroom the "discourse" was based on the Socratic method with the teacher asking many questions (more than two dozen) during her presentation, and calling upon students to answer, mostly from those who volunteered, but not entirely. There were

4.3 times as many males as females in this classroom, but they answered only 3.8 times as often, so the gender response ratio favored the females. However, because the females were so few in number in this classroom, despite the favorable ratio, it seemed the female students had little voice. There were a few instances during which the instructor had students working alone, in pairs (using think-pair-share), and as a whole group. The instructors' method of classroom management was effective in keeping all students on task. The instructions and directions were clear, and the entire lesson was well organized. The exit slip assessment used by this instructor was a three-question mini-quiz that included a written prompt asking students to reflect on what they had learned in class that day. Regarding physical manifestations of SV, student work and motivational sayings were posted on the board and walls. There were no physical manifestations of RM.

Instructor #2 Clinical Interview

When asked how important it was to her to create an equitable classroom I#2 replied, "I think it is important," and stated, "An equitable classroom provides an equal chance to participate in class. This may require something different or more encouragement for girls in class." She cited the women and mathematics course as influencing her desire for equity. She valued the course for making her aware of female mathematics role models for her students. I#2 mentioned two other courses that influenced her to work towards gender equity: a special education course taken as part of the credential program, and a course on coaching/mentoring taken while pursuing a graduate degree. As factors hindering her ability to achieve equity she mentioned two types of time limitations: finding time within the constraints of the curriculum to do more with women in math, and finding time to prepare special materials or lessons on women and mathematics. When asked if she felt she had been successful in creating a gender equitable classroom she replied, "I don't think so." She stated that she felt she needed more pedagogical tools/skills related to equitable discourse at her disposal.

Instructor #3 Classroom Observation

In her classroom, I#3 employed teaching methods that included a constructivist approach and some lecture using the Socratic method. The class opened with a hands-on guided discovery activity leading students to conjecture that the sum of the angles in a triangle is the same as a straight line. I#3 typically addressed questions to the entire class. Sometimes multiple individuals answered at once. While females answered the majority of the questions, the female to male gender response ratio of 1.2 to 1 fell well short of the 3 to 1 female to male ratio in the class that day. The small group work included a round-robin task that had students move in randomly assigned groups of three or four from one station to another to do problems. Due to effective classroom management, transitions from one mode of work to another were seamless. For the most part, all students seemed engaged and on task. The exception was during a round-robin activity when some students seemed disengaged while they waited for others in their group to catch up on the task. Again, the instructor made use of an exit slip to check understanding of the material presented along with a question about what they learned. Student work and motivational sayings posted on the walls comprised the physical manifestations of the SV. RM appeared in a single poster from the National Women's History Project (NWHP) featured a group of women mathematicians.

Instructor #3 Clinical Interview

I#3 stated that equity was very important to her and stated, "An equitable classroom requires pushing girls more toward participation and confidence." She cited the women and mathematics course as influencing her to strive for gender equity because it made her aware of female mathematics role models for both her students and herself. I#3 said she has made repeated reference to the women she learned about in the course and has used materials/lessons from the course with her own students on multiple occasions. I#3 noted two other influences that led her to strive for an equitable classroom. She described how she felt safe to contribute ideas in her senior seminar in mathematics because of how the instructor treated the students. According to her, he "gave equal attention to the males and females" and she felt "warmth in that class." He served as a powerful role model for equitable teaching, one she felt was difficult to come by in upper division math courses. The second experience had to do with a feeling of "belonging" during an undergraduate mathematics summer research experience. I#3 described how she had focused her classroom efforts on "pushing her girls." She was not fully satisfied with her achievement of an equitable classroom climate. She felt that she had been successful with the girls in one of her class periods (not the one observed), but was concerned that her strong emphasis on the girls' participation with statements like, "Come on girls, don't let me down!" might not be the best method to produce the results she wanted (equity for all). Next year she vowed to extend her equity focus beyond just girls to all her students, especially her English language learners and low math ability students.

Instructor #4 Classroom Observation

I#4 conducted the most student-centered and egalitarian discourse. She modeled her pedagogical approach after her mentor in the education department, the math methods teacher. Questions were directed to all students, a single student was never called on. Hand signals or answers written on small individual white boards were used to elicit participation and check understanding. Students were frequently directed to discuss a question or problem with a neighbor or to tell the meaning of a term to a neighbor, and students did so. I#4 seemed to be the most successful in achieving the SV equity markers of participation, confidence and engagement. The instructor circulated, answered questions and offered encouragement such as "You're on the right track." If she observed students not providing justification appropriately, she stopped to model for the whole class how to do so. A gallery walk was another method she employed to elicit participation among all students. This method entailed students moving about different stations (individually, in pairs, or small groups) to work problems posted on the walls. In I#4's class, there was no gender response ratio to be counted due to the nature of the discourse. It is worth noting that this instructor had more contact with the math methods teacher than the others. More specifically, she participated in a professional development program for in-service teachers, directed by the math methods teacher, which reinforced the student-centered discourse observed in her classroom. She too used an exit slip to assess comprehension. Of the three secondary instructors, I#4 had the most interesting student work (e.g., student "mind maps" for logic and reasoning and student depictions of Zeno's paradox) and motivational materials (college pennants and posters in addition to motivational sayings) on display in the classroom. She had the same women and math poster from NWHP as I#3, but she also had posters of specific male (Newton and Einstein) and female (Sonya Kovalevskaya and Sophie Germaine) mathematicians and posters showing applications of mathematics.

Instructor #4 Clinical Interview

For this instructor, having an equitable classroom is very important. She described it as, "100% of students in class engaged, not just listening but participating. The focus is on individual needs, so it does not mean equal attention to each student." Influencing her to strive for equitable practice, I#4 cited the women and mathematics course for making her aware of mathematical role models for her female students. The statistics presented in the course on women's participation and achievement in mathematics had also raised her awareness of gender inequities in mathematics. She said she has made repeated reference to several of the women she learned about in the course and has used materials/lessons from the course in her own classes on multiple occasions. She gave credit to her math methods course for helping her develop equitable teaching methods. She mentioned that it was a struggle to find time within the constraints of the curriculum to do more with women in math, but followed up with a remark that she was confident she could figure out a way to do it. Of the three secondary instructors, she seemed the most positive about her current level of success in achieving an equitable classroom.

Conclusions

All four instructors' interpretation of the meaning of equity matches the position taken by NCTM, namely that equity and equitable instruction requires them to do "something different" depending on what they perceive to be individual or group needs.

In regard to gender equity, there is alignment between how important they say it is to them, how they characterize it in terms of student voice and role models, and what they do, strive to do, or wish they knew how to do to achieve it. To elaborate, the instructors said equity is important to them, and that it is characterized by having an equal opportunity to learn that is taken advantage of by both males and females. For them, markers of gender equity will show up in the SV, where both genders participate, have confidence in their ability, and engage with the work. What the instructors do, try to do, or wish they were better equipped to do is to promote the SV of all. Each instructor identified areas where they felt they were or were not meeting their own standards and expressed frustration about the latter.

What we observed the instructors doing or having available in their classroom largely confirmed that their self-assessments are accurate. I#1, the first grade teacher, was generally satisfied with what she was achieving in her classroom in terms of SV and not too concerned about RM given the age of her students. Certainly, her teaching methods (whole group recitation, use of hands-on manipulatives, individual and small group work, Socratic-type discourse with the whole class, and incorporating reading into the math lesson) appropriately supported the SV of her young students. While her gender response ratio favored the males, the actual gender count of the responses was only off by one from the gender count in her classroom. Of the three secondary instructors, I#2 seemed the least positive about her current level of success in achieving an equitable classroom. Her classroom "discourse" was the most traditional, primarily relying on the Socratic method, with a small amount of think-pair-share. During this class period the gender response ratio favored the females, but there were so few females, it seemed they had little voice. The student work on display showed computations of area and perimeter for floor plans generated by the students but did not provide insights into student thinking. There were no role models for doing mathematics on display. Thus, it seems this instructor gave a realistic assessment of her own situation relative to SV and RM. I#3 was not fully satisfied with her current level of achievement with regards to an equitable classroom climate. Although females answered the majority of the questions asked during the observation, the gender response ratio favored the males in this predominantly female classroom. There was one RM poster for women in mathematics on the wall. In one class (not the one observed) her exhortations to the girls to be more like the boys produced positive results but she was uncertain about the effects of her methods on other groups (e.g., boys, English language learners, etc.). I#4 was the most positive in her assessment of her success in achieving an equitable classroom. The classroom observation certainly supported her assessment as she had achieved the most equity in SV through her pedagogical methods that elicited participation, confidence, and engagement among all students through hand signals, white boards, and pair and group discussions. Her classroom also had the most SV and RM material on display.

The instructors described a number of positive influences that promoted and supported their desire to achieve equitable classrooms. The women and mathematics class was successful in raising awareness of gender equity for all four instructors, either by providing access to role models or to statistical data about inequities. The course also gave them knowledge and materials they can use in their teaching (two of the four were actively using these materials). In addition to being in the same women and mathematics course, all four instructors had also taken the secondary mathematics methods course together. The influence of the math methods course on aspects of their teaching was confirmed by the observations. The fact that they adopted many of the strategies introduced to them by their math methods teacher (e.g., exit slips) provides strong evidence that the math methods course can play a key role in supporting equitable teaching. Finally, while three of the instructors did not identify any role models for equitable teaching practice, I#3 discussed the significance of having found one such role model in an upper division math content course.

Regarding hindrances to achieving an equitable classroom, the most cited factor was time: finding time to insert material on women and mathematics, and finding time to prepare special materials. One instructor wished she had more effective pedagogical tools at her disposal and another thought she needed to tone down rhetoric aimed at "pushing her girls" or she would be labeled 'sexist' and that could interfere with her rapport with her students.

Limitations of the Study

Whether one can or should draw broad implications for practice from case studies based on close examination of an unusual course, such as the course on women and mathematics, is questionable. This study was limited by the fact that only a single lesson was observed for each instructor. These lessons were not video-taped and therefore no analysis could be conducted of the types of questions asked of each gender, analysis commonly found in studies of gender equity and classroom discourse. Factors cited as influencing the instructors' views or practice were self-reports. The interview was conducted by the former professor from the women and mathematics course and that may have influenced the instructors' answers. Finally, due to time constraints, the coding was done by a single person. Still, common themes emerged from the interviews.

Implications for Practice

A number of implications for practice can be drawn. However, given the particular nature of this case study, it seems most appropriate to phrase these implications as reflective questions for college faculty and their programs:

- How well are our teacher preparation programs helping future K-12 teachers develop the necessary skills to achieve equitable classrooms?
- Is gender equity or diversity addressed in our math or teacher prep curricula, or coordinated in any way between them? Should it be? If so, how? By whom?
- How is "Student Voice" experienced in our own mathematics classrooms and in classrooms across our departments? Are math faculty encouraging participation, promoting confidence and achieving engagement among all students? How can equitable practice be highlighted and made more explicit in the women and mathematics course?
- Are role models available for *all* students in our collegiate mathematics departments? If they are not among our faculty, can they be found in posters in our hallways and common spaces, in our invited colloquia speakers, or in our career panels?
- Are gender and diversity concerns in STEM fields discussed in our departments or on our campuses? Do these discussions address both theory and practice? Do they engender change or action?

- What is the value in having collegiate math faculty visit the K-12 classrooms of former students and converse with them about their practice? Could those observations and conversations lead faculty to reflect on how to improve teacher preparation programs or encourage them to consider the environment in their own classrooms? If so, how?

References

Bonetta, L. (2010). Reaching gender equity in science: The importance of role models and mentors. *Science Careers*. Retrieved from http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2010_02_12/science.opms.r1000084

Dewar, J. (2008). What is mathematics: Student and faculty views. *Electronic Proceedings for the Eleventh Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education, San Diego, CA*. Retrieved from http://mathed.asu.edu/crume2008/Proceedings/Proceedings.html

Glaser, B. (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Halpern, D., Benbow, C., Geary, D., Gur, R., Hyde, J. & Gernsbacher, M. (2007, Fall). Why do men dominate the fields of science, engineering and mathematics? *Scientific American Mind.* Retrieved from http://www.sciam.com/article.cfm?id=sex-math-and-scientific-achievement

Hyde, J., Lindberg, S., Linn, M., Ellis, A. & Williams, C. (2008, July). Gender similarities characterize math performance. *Science. 321*(5888), 494-495.

Leonard, J. (2008). Women in aviation and space: The importance of gender role models in mathematics education. Chapter 6 (pp. 107-126) in *Culturally specific pedagogy in the mathematics classroom*. New York: Routledge.

Marx, D. & Roman, J. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*. *28*, 1183-1193.

National Coalition for Equity in Education. (2003). Perspectives on equity. Retrieved from http://ncee.education.ucsb.edu/perspectives.htm

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

National Council of Teachers of Mathematics. (2001). *Changing the faces of mathematics* (Series). Reston, VA: NCTM.

Perl, T. (1978). *Math equals.* New York, NY: Addison-Wesley.

Secada, W., Jacobs, J., Becker, J., & Gilmer, G. (2001). *Changing the faces of mathematics: Perspectives on gender*. Reston, VA: NCTM.

Steele, C., Spencer, S. & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in Experimental Social Psychology Vol. 37.* San Diego, CA: Academic Press.

Streitmatter, J. (1994). *Toward gender equity in the classroom: Everyday teachers' beliefs and practices.* Albany: State University of New York.

STUDENT REASONING ABOUT LINEAR TRANSFORMATIONS

Jessica Ellis, Frances Henderson, and Chris Rasmussen
San Diego State University

Michelle Zandieh
Arizona State University

In this report we detail linear algebra students' interpretations of linear transformations. Data for this analysis comes from mid semester, semi-structured problem solving interviews with 13 undergraduate students in linear algebra. We identified two main strategies used by students: 1) students used structural reasoning with entries of the matrix, columns of the matrix, and orientation of the shape and 2) students used operational reasoning through matrix and vector multiplication. We examine the patterns that emerged from student strategies, and discuss possible explanations for these patterns.

Key words: linear algebra, linear transformations, operational and structural reasoning, concept development

Introduction

A longstanding concern in mathematics education is the balance and relationship between two related types of knowledge: the knowledge needed to carry out computations and procedures, and the knowledge needed to understand the ideas and reasons that underlie these computations and procedures. The research community has addressed this issue through a number of explanatory frames, including procedural versus conceptual understanding (Hiebert, 1986), process versus object conceptions (Breidenbach et al., 1992), instrumental versus relational understanding (Skemp, 1976), synthetic versus analytic thinking (Sierpinska, 2000), and operational versus structural reasoning (Sfard, 1991). Although there are differences in these constructs (both nominal and theoretical), there is a general consensus that both types of knowledge are necessary in order to develop mathematical proficiency. Indeed, each of these types of understanding is represented in NCTM's five strands of mathematical proficiency (Kilpatrick, Swafford, & Findell, 2001), which highlights the need for students to develop both conceptual understanding and procedural fluency.

In the domain of linear algebra, researchers have expanded on these dual modes of reasoning. For example, Sierpinska (2000) describes different modes of student reasoning as *synthetic-geometric, analytic-arithmetic,* and *analytic-structural.* She notes that these modes of reasoning are neither mutually exclusive nor hierarchical, but that the synthetic mode represents "the practical way of thinking" and the analytic represents the "theoretical way of thinking" (p. 233). Related to these modes of reasoning, Hillel (2000) describes three modes of representations: *geometric* (using the language of R^2 and R^3, such as line segments and planes), *algebraic* (using language specific to R^n, such as matrices and rank), and *abstract* (using the language of the general formalized theory such as vector spaces and dimension).

A number of student difficulties in linear algebra have also been documented (see Carlson, 1993; Hillel, 2000; Dorier, Robert, Robinet, & Rogalski, 2000; Sierpinska, 2000; Stewart & Thomas, 2009), with many of these difficulties attributed to the disconnect between various representations and students' modes of reasoning. For example, some researchers have been interested in how a geometric introduction to linear algebra may (or may not) help students make connections to algebraic and abstract modes of reasoning. As part of a teaching experiment involving a geometric introduction to linear transformations, Sierpinska et al. (1999) found that students' difficulties in linear algebra persisted even when provided concrete and visual connections to the theory. Sierpinska suggests that one reason for students' difficulties may be that the content, while geometric, emphasized purely analytic understanding rather than involving synthetic understanding as well. These findings provide further evidence for Carlson's (1993) suggestion that students are most naturally proficient and comfortable with computational approaches (i.e. relying on synthetic reasoning), and lead Sierpinska to question whether we should encourage synthetic reasoning in addition to analytic reasoning *throughout* students' learning. This suggestion contrasts Sfard's (1991) belief that concept development occurs as one transitions from operational to structural understanding, and raises the questions of what role these various modes of reasoning play in students' understanding of linear algebra and when these modes of reasoning are engaged by students in relation to their level of 'concept development.'

In this study we seek to explore these questions by examining students' conceptions of linear transformations through an exploration of their solutions to a series of tasks involving geometric representations of linear transformations. These tasks differed in their level of complexity: in increasing order of complexity, the first task was a matching problem, the second was a prediction problem, and the third was a creation problem. The research questions related to these tasks are:

1. What are students' strategies on these types of problems?
2. What patterns exist in students' strategies across the three types of problems?

In answering these questions we also sought to account for any patterns identified in student reasoning.

Methods

Data for this analysis were collected from one extensive, semi-structured problem-solving interview (Bernard, 1988) with 13 undergraduate students. The interview questions were used to gather information related to participants' understanding of linear transformations, with an emphasis on geometric representations of linear transformations. For this study, the last three questions of the interview were analyzed: a matching question consisting of five parts, a prediction task, and a creation task. These tasks will be discussed in detail below. The students came from a large southwestern university and were primarily engineering majors. Four of these students received a final grade of a 'C' in the linear algebra course, six students received a 'B', and three received an 'A'. Pseudonyms were developed that reflect these grades. The interview was the second of a series of three that was part of a semester-long classroom teaching experiment (Cobb, 2000). The interview was conducted after students had discussed geometric and algebraic interpretations of linear transformations, but before they had begun a unit on eigen-theory.

Each interview was videotaped, transcribed, and thick descriptions were developed for students' solutions to each of the tasks that included students' written work (Geertz, 1994). The videos, transcriptions, and thick descriptions were analyzed through grounded analysis (Corbin & Strauss, 2008). Two of the researchers independently watched each of the 13 students' responses to the first three matching tasks in order to begin analysis. These researchers determined that there was a clear difference between students' strategies that could be differentiated as treating the matrix as a process or treating it as an object. Keeping this distinction in mind, the researchers then watched the last two matching tasks for a subset of seven students and realized that there were three different ways students treated the matrix as a process and three ways that students treated the matrix as an object. Because linear transformations are functions, we chose to employ Sfard's (1991) distinction between operational and structuring reasoning as a framework for classifying six a posteriori strategies.

Two researchers then watched the remaining six students' solutions to the last two matching tasks and applied the emergent coding, noting each strategy used in the order that it was used, resulting in an overall strategy for each task. The researchers compared their coded strategies for each of the six students, and there was agreement on the vast majority of coding. Any discrepancies were discusses and final codes were agreed upon. The remaining tasks (the first three matching tasks, the prediction task, and the creation task) were then coded for all students; one researcher coded six students and the other coded seven students. Throughout this final coding process we developed operational definitions for each of the six categories, refining these definitions as we deemed fit, which will be discussed in the results section below.

Interview Tasks

We analyzed student responses to three tasks from the interview: a matching task, a prediction task, and a creation task. The matching task consisted of five problems of increasing difficulty, beginning with a positive, diagonal matrix and ending with a matrix with all non-zero entries. The prediction task was created to be slightly more difficult than the matching tasks, and the creation task was thought to be the hardest. This task design was modeled after Artigue's (1992) interview task design involving student understanding of differential equations.

The prompt for each matching task was as follows: "In each of the following questions, you are given a matrix transformation and a corresponding set of images. Identify any images that correspond to the image of the unit square (as shown below on the left) under the given transformation." There were five parts, each part involving a different matrix and a different set of possible images under the given transformation. The five matrices that were provided were: $A = \begin{bmatrix} 2 & 0 \\ 0 & 1/3 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 0 & -1/3 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 2 \\ 1/3 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$, and $E = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. See *Figure 1* for the first of the five matching tasks. The prediction task asked students to: "Please find the image of the picture below under the matrix transformation $\begin{bmatrix} 3 & 1 \\ 0 & -1 \end{bmatrix}$" and provided an image of a 'T', as in *Figure 2.* The creation task required students to find a matrix that that fit a given transormation, as represented by an initial and final figure. The prompt was: "Please find a matrix that transforms the image on the left into the image on the right. Note that the rectangle on the left is 3 units by 2 units." Students were shown two images of a 3 x 2 rectangle, one 'untransformed' and one transformed under a to-be-determined matrix, as shown in *Figure 3*.

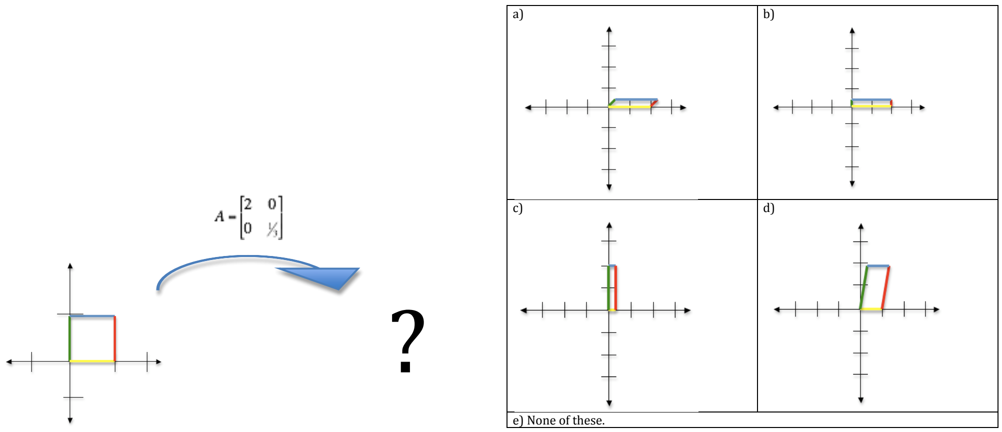


Figure 1. Task A of five matching tasks.

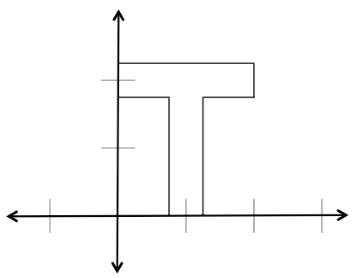


Figure 2. Prediction task.

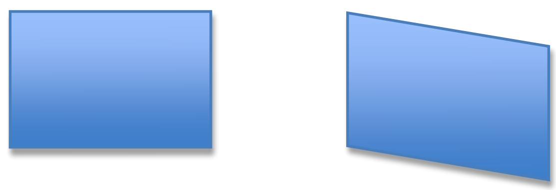


Figure 3. Creation task.

Results

In this section we present analysis of students' strategies while solving the matching, prediction, and creation tasks. Students approached these tasks with a wide variety of strategies, and appeared to either view the matrix as a tool that performs the actions of the transformation (for example, by inputting vectors into the matrix to compute the resultant vector), or as an entity that provides information about how the transformation acts (for example, what do the individual entries in the matrix tell you, or what do the columns of the matrix tell you). We interpreted these different conceptions as viewing the matrix as a process or viewing it as an object, and made use of Sfard's (1991) distinctions between *operational* and *structural* conceptions to differentiate students' solutions.

Student reasoning on these tasks were further classified into six strategies, three of which related to a structural conception of linear transformation and three to an operational conception. We refer to these six strategies as *Structural entries (S_e)*, *Structural vector (S_v)*, *Structural orientation (S_o)*, *Operational identify (O_i)*, *Operational unit-vector (O_u)*, and *Operational vector (O_v)*. We operationally defined each of these categories as follows:

Structural entries (S_e) – A student categorized as using an S_e strategy reasoned by treating the two by two matrix as being composed of four pieces, the entries of the matrix. These students would reason using the entries as indicators of how the box stretched or shrunk along a particular axis, slanting or shearing of the box, as well as flipping or rotating of the unit box.

Structural vector (S_v) – A student categorized as using an S_v strategy reasoned by treating the two by two matrix as being composed of two pieces: the two column vectors of the matrix.

These students reasoned about the column vectors of the given matrix as the sides of the transformed box without doing any computations or reasoning with the entries.

Structural orientation (S_o)- A student categorized as using an S_o strategy attended to the visual and/or geometric properties of the original shape/graph as opposed to properties of the matrix. *So* often appeared when the students discussed the orientation of the box as well as how the colors of the sides should be oriented.

Operational identify (O_i)- A student categorized as using an O_i strategy reasoned by performing multiplication with the identity matrix. These students carried out their computation using the identity matrix, which we posit represented the unit box.

Operational unit-vector (O_u)- A student categorized as using an O_u strategy reasoned by performing multiplication dealing with the unit vectors. In the matching tasks, the unit vector (1,0) was colored green, and the unit vector (0,1) was colored yellow, and thus students who performed operations on the 'green' and 'yellow' vectors were considered to be employing this strategy.

Operational vector (O_v)- A student categorized as using an O_v strategy reasoned by performing multiplication dealing with a non-unit vector, such as (1,1).

Next we provide illustrative examples for each strategy:

Structural entries (S_e): In the following excerpt, Becca determined what the unit square would look like under the transformation A by performing computations of vectors. After she obtained her answer, the interviewer asked if she had an intuitive way of knowing what the answer would be without doing the computations. Becca's answer was categorized as *Se* based on her explanation of how the non-zero entries of the matrix visually affective the unit square:

Interviewer: *Is there a way, do you have any way to look at your choice and look at the matrix and be like, 'yeah, that kind of makes sense?'*

Becca: *...this is 2 [underlines the 2 in the matrix], and that's 1/3 [underlines 1/3], so that's stretching it 2 in the x direction and then I guess shrinking it to 1/3 in the y direction.*

Structural vector (S_v): In the following excerpt, Caden determine what the unit square would like look under the transformation D without writing anything down. When asked how he determined the (correct) answer, Caden explained that he saw the transformation as a box made out of the column vectors of the matrix D.

Interviewer: *Can you tell me some more how you came to [your answer]?*

Caden: *Well, that's just these vectors (marks the column vectors of D* $D=\begin{bmatrix}1&0\\1&-1\end{bmatrix}$ *). Made into a box (marks the vectors on (c)*  *). Multiplied by the unit vector, yeah, it seems logical.*

Structural orientation (S_o): In the following excerpt, Ben determined what the unit square would like look under the transformation B without saying or writing anything and seemed unsure about his answer. The interviewer asked him about how he arrived at his answer, and "what his inner debate was" and Ben's response indicates that he reasoned about the problem initially without the colors, but then realized that the colors played a significant role in the orientation of the unit square. He then obtained his answer by reasoning with both the colors as well as the visual changes in the orientation of the figure under the transformation:

Ben:	*Um, well. At first I wasn't really paying attention to the colors of the lines, and so for a second...these two [possible answers] (a and c), look to be right. And then I had to stop and pay attention to the colors. So I think it's flipping it over, and whatever shift in size is happening. So I think it's (c), since I think you're changing the shape and then flipping it over the x axis. Since it's going from 0,1 to 0,-1/3, yeah.*

Operational identify (O_i): In the following excerpt, Chad began to determined what the unit square would like look under the transformation D by using the entries of the matrix, but became unsure if the 1 in the off diagonal would result in a skew. When asked if he had a way to check, Chad performed a calculation using the identity matrix. When asked what this calculation told him, Chad replied that it showed him "the new square that we're looking for," but was unable to make further progress with the problem:

Interviewer:	*Do you have a way that you could check?*
Chad:	*I guess if I multiply the matrices (writes* $\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}1&0\\1&-1\end{bmatrix}$*).*
Interviewer:	*Can you explain to me why you're, here I see you've gotten D?*
Chad:	*Yeah, this one [points to the identity matrix] is that, the unit square, so you multiply them.* $\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}1&0\\1&-1\end{bmatrix}=\begin{bmatrix}1&0\\1&-1\end{bmatrix}$
Interviewer:	*So you saying if you multiply by the unit square, what is that going to tell you?*
Chad:	*That would be the transformation.*
Interviewer:	*Which would be the transformation?*
Chad:	*We're taking the unit square and we're multiplying it by this matrix [points to matrix D].*
Interviewer:	*Then what is this [points to the result after the multiplication]?*
Chad:	*That should be what the new one is [points to the result after the multiplication]?*
Interviewer:	*The new?*
Chad:	*The new square that we're looking for.*

Operational unit-vector (O_u): In the following excerpt, Becca determined what the unit square would like look under the transformation D by immediately using computations involving the unit vectors (the green and yellow vectors) and successfully found the matching transformed unit square through these computations:

Becca:	*1,0; 1,-1. So I'll just compute it. This is yellow. So green. 0, -1.*

$$\text{yell} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\text{green} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

So yellow should be 1,1 like that [points to (c)]. And then green should be 0,-1, so (c).

Operational vector (O_v): In the following excerpt, Bart first notes that he can't determine what the unit square would like look under the transformation E by "just how the matrix is set up" so he computes where the matrix E would send the vector (1,1) to narrow down the possible answers:

Bart: *This one I won't be able to rule out by. I just can't think about it that way; I can't rule it out by just how the matrix is set up. I know it stretches by 1 in the x, -1 in the y, skews by 1 in the x and 1 in the y, but I just can't picture how that would work with the unit square. So I'll have to do some computations.*

[writes $E = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ *] Times 1,1 is equal to 2, 0. So that vector changes to this vector. So it's none of the others, it has to be this one [c]. If anything at all, so it's either (c) or (e).*

Frequently students' overall strategies for solving these tasks involved many sub-strategies; for example a student may solve a task by using an overall strategy of $S_eO_uO_v$ (first using the entries of the matrix, then performing computations on both unit vectors and non unit vectors). In Table 1 below we report students' overall strategies for each task. Sub-strategies were coded in order of use. For example, on matching task d, Alex used an overall strategy of $S_eO_vS_o$, indicating that he first used the entries of the matrix to inform his solution, then performed a computation using a non-unit vector, and last reasoned about the orientation or colors of the matrix. When a student switched from a structural strategy to an operational one (as in S_eO_v), often the student tried to solve the task structurally and was unable to, so chose to employ an operational strategy. When a student switched from an operational strategy to an operational one (as in O_vS_o), often the student found their answer operationally, and then either checked their answer structurally or connected their answer back to the originally matrix by reasoning structurally. Entries that are highlighted in blue indicate that these strategies relied only on structural strategies, and those highlighted green indicate that a purely operational strategy was employed.

This table was the main data source used for the analysis of these tasks. We also developed Table 2, which shows predominant student strategies grouped by task and students' final grades (the usage frequency is noted in parentheses). These tasks were grouped as follows: the matching tasks into three groups (the diagonal matrices (a and b), the non-diagonal matrices with at least one zero entry (c and d), and the matrix with no zero entries (e). Finally, Table 3 was created to show the average time spent by each grade group on each group of tasks.

The analysis of the data was conducted in two ways: first we looked for patterns *within* each of the individual tasks, and then we looked at the individual student strategies *across* the tasks.

Table 1. Student reasoning by task

	Match. a	**Match. b**	**Match. c**	**Match. d**	**Match. e**	**Prediction**	**Creation**
Alex	S_e	S_eS_o	$S_eO_vO_u$	O_vS_o	$S_eS_oO_v$	$O_vS_oO_v$	O_vS_e
Andrew	S_eO_u	S_e	S_e S_vS_o	$O_uO_vS_o$	$O_uO_vS_o$	$S_eO_vS_o$	S_e
Anthony	S_e	$S_eS_oO_v$	$S_eS_oO_v$	$O_vO_uS_o$	$O_vS_oS_v$	$O_vS_oS_e$	S_eO_v
Bailey	$O_uO_vS_e$	S_e S_o	O_u	O_u	O_uO_v	$S_oO_vS_e$	O_vS_v
Bart	$S_vS_eO_v$	S_e S_oS_v	$O_vO_uS_o$	$S_eO_vO_u$	$S_eO_vO_u$	O_vS_o	O_vS_o
Becca	O_uS_e	$S_eO_uS_o$	$S_eS_oO_uS_v$	$O_uS_vS_oO_v$	O_uO_v	$O_vS_eO_vS_o$	O_vS_e
Ben	O_u	O_uS_o	O_uS_o	O_uS_o	O_uS_o	$O_vS_eS_o$	O_v
Bill	S_eO_u	O_uS_o	$S_eO_uS_oO_v$	O_u	O_u	$S_oO_vS_e$	O_v
Brad	$S_vO_iS_e$	S_e	$O_iS_vS_o$	O_iS_e	S_o	S_eS_o	S_eO_v
Caden	O_iS_o	S_oO_v	S_e	S_vS_o	S_vS_o	S_eS_o	S_vS_o
Chad	S_e	S_e S_o	SeS_o	S_eO_i	$S_oO_iS_v$	$O_vS_eS_o$	S_eO_v
Charles	S_eO_i	S_e S_o	S_e	$S_eS_oO_iS_v$	S_oS_v	$S_eS_oO_v$	$S_eS_vO_v$
Chris	S_e	S_e S_o	$S_eO_vO_uS_o$	S_eS_o	S_eS_o	S_eS_o	S_eS_o

Table 2. Predominant student reasoning by grade and overall

	Match. a and Match. b	**Match. c and Match. d**	**Match. e**	**Prediction**	**Creation**
A	S_e (6/6)	S_e/ O_vS_o (4/6)	O_vS_o (3/3)	O_vS_o (3/3)	S_e, O_vS_e (3/3, 2/3)
B	S_e/O_u (8/12, 7/12)	S_oO_u (10/12)	S_oO_u (4/6)	$O_vS_oS_e$ (4/6)	O_v (6/6)
C	S_e (6/8)	S_eS_o (4/8)	S_v/S_o (3/4)	$O_vS_oS_e$ (2/4)	S_e (4/4)
Overall	S_e (20/26)	S_o/ S_oO_u (18/26, 13/26)	S_o (10/13)	O_vS_o (9/13)	S_e (8/13)

Table 3. Average time on task by grade and overall

	Match. a and Match. b	**Match. c and Match. d**	**Match. e**	**Prediction**	**Creation**
A	1:56	7:17	2:05	10:53	17:04
B	3:03	4:52	1:50	11:26	13:04
C	2:55	5:25	7:29	10:28	10:33
Overall	2:38	5:51	3:48	10:56	13:34

Within Task Analysis

Overall, we see that students predominantly used a combination of operational and structural strategies. However, on the first two matching tasks (the diagonal matrices), eleven students used a strategy of *only* structural reasoning. On the diagonal matching tasks, each grade group of students used Se as their most common strategy. Interestingly, the B-students also predominantly used Ou (7 out of 12 times) while the A-students used this strategy only once (out of 6 times) and the C-students never used Ou (out of 8 times). Overall for these tasks, Se was used 20 out of 24 times.

On the non-diagonal tasks that had at least one zero, we see more use of operational strategies: A-students used both Se and SoOv most often, B-students used SoOu most often, and C-students used SeSo most frequently. Overall, the most predominant strategy was So (used 18 out of 24 times), and a combination of So and Ou was used on over half of the attempts (13 out of 24). These tasks also proved to be the most difficult for students: four out of six attempts from the A-students were partially or entirely incorrect, half of the B-student attempts were at least partially incorrect, and six out of eight of the C-student attempts were at least partially incorrect.

On the last matching task, arguably the most difficult, only two out of the thirteen solutions were at least partially incorrect. The A-students all employed a strategy involving OvSo, the B-students predominantly used a mixed strategy of SoOu, and the C-students favored strategies involving both Sv and So.

On the prediction task, we see that the predominant strategy for each grade group involves a combination of both structural and operational reasoning: the A-students all used a strategy involving OvSo, the B-students used a combination of SoOu, and the C-students used either Se or So predominantly. Interestingly, only one of the four C-students employed a strategy involving an Operational method, whereas every A-student and all but one B-student used a mixed strategy. Across the board, students performed well on this task.

On the creation task, student strategies varied significantly by grade-group. The A-students preferred mixed strategies, while the B-students favored operational methods and the C-students favored structural methods. Overall, Se was the most predominantly used strategy. As we expected, this task appeared to be more difficult than the others based on the length of time it took students to complete and the two incorrect solutions.

Across Task Analysis

One of the most striking patterns that we recognize is the overall shift from predominantly structural strategies to a combination of operational and structural strategies. On the first two tasks, students employed a strategy involving only structural reasoning eleven out of the total 26 times (42% of the time), whereas on the remaining tasks a purely structural method was used 16 out of 65 times (24% of the time). Overall, the C-students employed purely structural strategies overwhelmingly more than the other students. If we focus only on the A and B students, this trend becomes even clearer: on the first two matching tasks, these students employed a purely structural method seven out of 18 times (39% of the time); on the remaining tasks a purely structural strategy was used only four out of 45 times (9% of the time). We also see a surprising absence of operational strategies: overall, a purely operational strategy was employed only five times out of 91 times (6% of the time), whereas a purely structural method was employed a total of 27 times (30% of the time). However, the B-students were the only group of students to employ a purely operational strategy: the A-students and the C-students *never* employed an operational only strategy. We see a slight difference between correctness of solutions based on students' grades: A-students were correct 76% of the time, B-students 81% of the time, and C-students 68% of the time.

Discussion

Our goal in this discussion is to examine various explanations to account for the patterns discovered in the within and across task analyses. We have determined three possible explanations for these differences that will be discussed here: (1) the nature of the tasks; (2) concept development; (3) and knowing-to use certain strategies.

Task sequence

One surprising pattern was that as the tasks progressed in difficulty, students' did *not* do worse; on the most difficult task (the last matching task- from the researchers' perspective), the students were correct more frequently than on less difficult tasks (the matching tasks with the non-diagonal matrices and at least one zero entry). One explanation for this is that the sequence of tasks aided students' understanding of how to solve the tasks. We see that as students transitioned from reasoning about a diagonal matrix to a non-diagonal matrix, they struggled more. However, transitioning from a non-diagonal matrix with at least one zero entry to a matrix with no zero entries was less difficult for students: students improved on both correctness and efficiency. Once students progressed through the sequence of matching tasks, they appeared well prepared to reason about the prediction and creation tasks. These tasks, which were designed to be novel and cognitively demanding, appeared to be well understood by the majority of the students –as demonstrated by the correctness of their solutions. In the final task students were shown a 3x2 rectangle transformed under an unknown transformation and are asked to create the transformation resulting in the image shown. Only two of the 13 students did not correctly identify the linear transformation using a combination of strategies.

We suspect that had this task been shown to students first in the sequence instead of last, students would not have been as successful. This speculation would be very interesting to look into in further research. Additionally, the affects of this particular task sequence have implications in the teaching of linear algebra. If one of the goals of teaching linear algebra is for students to develop strong understanding of matrices both algebraically and geometrically, this sequence of tasks may prove beneficial. One of the researchers is teaching linear algebra again in the upcoming term, and is informally exploring the benefits of this task sequence in the classroom. Further and more formal studies of this task sequence in the classroom would be very interesting.

Concept development

One of the clearest patterns that we saw in the data was transition from predominantly structural to a combination of structural and operational reasoning. Sfard (1991) described concept development as a shift from an operational conception to a structural conception. Thus, we may explain this shift in student strategies as indicative of students' stronger understanding of the geometric implications of linear transformations represented by diagonal matrices versus transformations with matrix representations that contain non-zero entries on the non-diagonals, prediction tasks or creation tasks. This is not surprising, especially considering the geometric results of diagonal matrices versus non-diagonal matrices, and the visual ease of understanding stretching compared to skewing.

What is surprising is that C-students overall exhibited a much higher frequency of purely structural strategies. Do C-students have fuller concept development of the geometric implications of linear transformations than A and B-students? Or, do C-students have a weaker operational understanding of matrices and thus instead rely on their structural conceptions? In these tasks we were not specifically interested in how strong students' procedural competency was, and thus have no way to assess if this explains C-students' preference for structural strategies. However, a weak understanding of matrix multiplication certainly would result in a low grade in any linear algebra course. An alternate explanation for C-students' seemingly preferred use of structural strategies may be found in Sfard's (1991) construct of *pseudo-structural* reasoning: pseudo-structural strategies would be classified as those in which the student deals with the transformation represented by the matrix in a structural fashion, but cannot unpack their understanding and show the underlying operational understanding. These differences suggest that further investigation into the differences between A, B, and C-students' operational and structural conceptions is needed.

Knowing-to

In addition to the C-students' preference towards purely structural strategies, the B-students were the only students to employ purely operational strategies- and did so very frequently. Can we similarly question if B-students have a fuller operational understanding than the other students, or a weaker structural understanding? Instead of these questions, the more appropriate question may be: why do A-students prefer strategies that combine both structural and operational conceptions? One explanation for these differences may be found in Johnston-Wilder and Mason's (2004) construct *knowing-to.* Johnston-Wilder and Mason expand on Ryle's (1949, as cited in Johnston-Wilder and Mason, 2004) distinction between *knowing-that* (factual), *knowing-how* (to perform acts), and *knowing-why* (having stories to account for phenomena and actions), by adding *knowing-to* (acting in the moment as deemed appropriate). They note that educators' dismay that students who have shown that they know-how to carry out a procedure yet fail to do so on a test can be attributed to a lack of knowing-to apply this knowledge. One way to account for A-students' flexibility in transitioning between operational and structural strategies may be that the A-students in this study have a stronger sense of knowing-to apply the appropriate strategy at the appropriate time, whereas weaker students may tend to persist with a strategy that has proven useful. In order to test this conjecture, further examination of students' transitions between strategies will be done. Many would agree that A-students *should* have a stronger understanding of when to apply certain procedures or reason in other ways. This is a knowledge that should distinguish between A-students and B and C-students.

This paper has identified and described various strategies that students may employ when solving tasks involving geometric interpretations of linear transformations, and has related these strategies to operational and structural conceptions. We have discussed the patterns that emerged within the various tasks across the student grade groups, as well as across the tasks. We saw significant differences in the ways A, B, and C-students solved these tasks, and have explored various explanations for these differences. Further investigations into how students' course grades may or may not relate to their exhibition of operational and structural conceptions in linear algebra tasks and beyond would be a fruitful endeavor.

References

Artigue, M. (1992). Cognitive difficulties and teaching practices. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (pp. 109-132). Washington, DC: The Mathematical Association of America.

Bernard R. H. (1988). *Research methods in cultural anthropology*. London: Sage.

Breidenbach, D., Dubinsky, E., Hawkes, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics, 23,* 247-285.

Carlson, D. (1993). Teaching linear algebra: Must the fog always roll in? *The College*

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. 307-334). Mahwah, NJ: Lawrence Erlbaum Associates.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for doing grounded theory research*. Los Angeles: Sage Publications.

Dordrecht: Kluwer. 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. 177-189). Dordrecht: Kluwer.

Dorier, J-L., & Sierpinska, A., (2001) Research into the teaching and learning of linear algebra. In Holton, D. (Eds.), The Teaching and Learning of Mathematics at University Level: an ICMI study. Dortrecht, The Netherlands: Kluwer Academic Publishers, pp. 255-274

Dorier, J.L., Robert, A., Robinet, J., & Rogalski, M. (2000). The obstacles of formalism in linear algebra. In J.L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 85-124).

Geertz, C. (1994). Thick description: Toward an interpretive theory of culture. In M. Martin & L. C. McIntyre (Eds.), Readings in the philosophy of social science (pp. 213-232). Cambridge, MA: The MIT Press.

Hiebert, J. (1986). *Conceptual and procedural knowledge: The case of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Hillel, J. (2000). Modes of description and the problem of representation in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 191-207). Dordrecht: Kluwer Academic Publisher.

Johnston-Wilder, S. and Mason, J. (2004). Fundamental constructs in mathematics education. Routledge Falmer in association with the Open University.

Kilpatrick, J., Swafford, J., & Findell, B. [Eds.] (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Mason, J. (2002). Researching your own practice: the discipline of noticing. UK: Routledge

Mathematical Journal, 24(1), 29-40.

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.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), pp. 4-13

Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 209-246). Dordrecht: Kluwer Academic Publisher.

Sierpinska, A., Trgalova, J., Hillel, J., & Dreyfus, T., (1999). Teaching and Learning Linear Algebra with Cabri. Research Forum paper, in The Proceedings of PME 23, Haifa University, Israel, Volume 1, 119–134.
Stewart, S., & Thomas, M.O.J. (2009). A framework for mathematical thinking: the case of linear algebra. *International Journal of Mathematical Education in Science and Technology ,40*(7), 951-961.

UNDERSTANDING CALCULUS BEYOND COMPUTATION: A DESCRIPTIVE STUDY OF THE PARALLEL MEANINGS AND EXPECTATIONS OF TEACHERS AND USERS OF CALCULUS

Leann Ferguson
Indiana University, Bloomington

Calculus is an important tool for building mathematical models of the world around us and is thus used in a variety of disciplines, such as physics and engineering. These disciplines rely on calculus courses to provide the mathematical foundation needed for success in their courses. Unfortunately, due to the basal conceptions of what it means to understand calculus, many students leave their calculus course(s) with an understanding misaligned with what is needed in the follow-on discipline courses and are thus ill-prepared. By working with presumed experts (undergraduate mathematics and other discipline faculty members) to develop a small number of prototype tasks that elicit, document, and measure students' understanding of a few calculus concepts they believe are essential to successful academic pursuits within their respective disciplines, this study documents how the faculty participants' underlying conceptions about understanding changed and converged. Implications for calculus instruction and curriculum are mentioned.

Key words: Calculus, understanding, STEM preparation, design research

> *Mathematics can and should play an important role in the education of undergraduate students. In fact, few educators would dispute that students who can think mathematically and reason through problems are better able to face the challenges of careers in other disciplines – including those in non-scientific areas. Add to these skills the appropriate use of technology, the ability to model complex situations, and an understanding and appreciation of the specific mathematics appropriate to their chosen fields, and students are then equipped with powerful tools for the future.*
>
> *Unfortunately, many mathematics courses are not successful in achieving these goals. Students do not see the connections between mathematics and their chosen disciplines; instead, they leave mathematics courses with a set of skills that they are unable to apply in non-routine settings and whose importance to their future careers is not appreciated. Indeed, the mathematics many students are taught often is not the most relevant to their chosen fields. … The challenge, therefore, is to provide mathematical experiences that are true to the spirit of mathematics yet also relevant to students' futures in other fields.*
>
> (Ganter & Barker, 2004, p. 1)

These claims detail the rationale for The Mathematical Association of America's (MAA) Curriculum Foundations Project (CFP, http://www.maa.org/cupm/crafty/cf_project.html). Portions of the mathematics community and its partner disciplines, what I refer to as "client" disciplines (e.g., biology, business, chemistry, computer science, various areas of engineering), worked together to generate a set of recommendations that have assisted mathematics departments plan their programs to better serve the needs of client disciplines (Ferrini-Mundy & Gücler, 2009).

What does it mean for a mathematics course (e.g., calculus) to serve the needs of client disciplines? More often than not, client departments expect the pre-requisite calculus course(s) to provide the mathematical foundation needed for success in their calculus-based courses (Klingbeil, Mercer, Rattan, Raymer, & Reynolds, 2006). Are the calculus courses emphasizing the understanding needed for success in the client courses? Much research shows they are not and the graduates of the calculus course(s) leave with an "exceptionally primitive" understanding of fundamental calculus concepts (Ferrini-Mundy & Graham, 1991;

Zerr, 2010) and are ill-prepared for client courses (Ganter & Barker, 2004; Kasten, 1988; Klingbeil, et al., 2006).

As Ganter and Barker (2004) implied, client department faculty often complain that students are unable to apply calculus in the client coursework. Sometimes this coursework asks students to use the calculus concepts in ways not familiar to them. At other times, even when the concept is used in a similar fashion, differences in notation or a lack of familiar cues derails students. Such difficulties in transferring knowledge between disciplines are stark indicators of a lack of understanding (Hughes Hallett, 2000). Muddying the waters further are the numerous characterizations offered by literature that do not clarify what it means to understand calculus (Hiebert et al., 1997), much less provide resources for measuring this understanding.

Ferrini-Mundy and Gücler's (2009) review of the education reform efforts put forth within the undergraduate STEM disciplines provided an indication of the nation's willingness and commitment to ensure students learn these disciplines to the levels needed for competitiveness and for literacy. Before students can compete nationally, they must be successful within the academic world. Success in this world requires an applicable understanding of calculus because "modern scientific thought has been formed from the concepts of calculus and is meaningless outside this context" (Bressoud, 1992, p. 615).

The changes during the reform years placed greater emphasis on conceptual understanding (Hughes Hallett, 2000), but as Ganter and Barker (2004) pointed out, it has not been enough; the disconnect between what the client disciplines need and what the calculus courses provide still exists. For this reason, this study sought to answer these questions:

1. What are the different disciplines' perceptions of calculus?
2. What calculus concepts are needed and in what context(s)?
3. What does it mean to understand calculus?
4. How will teachers know if their students understand calculus?

Following in the footsteps of the CFP, this study explored the potential disconnect between the calculus taught and the calculus used at a particular undergraduate engineering institution. Through exploring this disconnect, this study identified several fundamental calculus concepts students need for successful academic pursuits outside the calculus classroom. This study pushed beyond the CFP by describing what it means to understand these concepts and developing tasks that allow teachers to assess calculus understanding.

Description of Study

Describing the fundamental calculus concepts and developing the prototype tasks constituted a design research study (Brown, 1992; Collins, 1992). As design research, each cycle of this study included divergent ways of thinking, selection criteria for the most useful ways of thinking, and sufficient means of carrying forward the ways of thinking so they could be tested during subsequent cycles.

Twenty-one faculty members (9 "teachers" and 12 "users"[1]) at an undergraduate engineering institution participated in an iterative series of interviews during which they expressed, tested, and revised the descriptions of fundamental calculus concepts, frameworks for understanding each concept, and draft tasks. Mathematics and client department faculty were selected based on their proximity to the calculus courses and the calculus-based client courses.

[1] For the purposes of this study, I define "teachers of calculus" as those faculty participants that have taught and/or are teaching Calculus I and Calculus II. "Users of calculus" are those faculty participants that have taught and/or are teaching the first client discipline course(s) that list Calculus I or Calculus II as a pre-requisite or co-requisite.

The framework for this study can be thought of as a multi-tier design experiment (Lesh & Kelly, 2000). As Figure 1 outlines, there were three tiers in this research project: 1) students, 2) faculty members/researchers, and 3) researcher/facilitator. For the research described here, the goal was not to produce generalizations about students or faculty members. Instead the primary goal was to work with presumed "experts" (instructors that taught a course of interest two or more times) to develop a small number of prototype tasks to elicit, document, and measure students' understanding of a few calculus concepts the faculty participants believe to be essential to successful academic pursuits within their respective disciplines.

Tier 3: The Researcher Level	Researcher develops models to make sense of faculty members' and students' calculus understanding. The researcher's interpretations are revealed through facilitation of the faculty interviews and student work sessions. Describing, explaining, and predicting faculty and student behaviors and responses further reveals the researcher's interpretations.
Tier 2: The Faculty Level	As faculty members develop shared tools (such as guidelines to assess student responses) and as they describe, explain, and predict students' responses, they construct and refine models to make sense of students' calculus understanding.
Tier 1: The Student Level	Individual students work on several tasks in which the goals include eliciting, documenting, and measuring the individual student's calculus understanding.

Figure 1. A three-tiered design experiment (adapted from Lesh & Kelly, 2000).

The calculus concepts (function, limit and continuity, rate of change, accumulation, and the fundamental theorem of calculus), together with frameworks and tasks, believed to be fundamental by mathematics and mathematics education researchers formed the basis of the interviews. After each cycle, a compare/contrast analysis was conducted on the emerging products. This consolidated set of products formed the basis of the next cycle. Subsequent cycles centered on analyzing and revising the tasks. Tasks were evaluated and analyzed first through the faculty lens[2] and then through the medium of student work. Task modification and writing completed each interview. The cycles, including data collected and products, are outlined in Figure 2.

Data Collection	Goals	Data Collected	Products
Cycle 1: Describing Calculus and its Fundamental Concepts, Developing Draft Tasks	▪ Make explicit what calculus is and how students need to understand the necessary calculus concepts within respective disciplines ▪ Develop drafts of prototype tasks	▪ Interview notes from each intradisciplinary group ▪ Draft tasks ▪ Audio and video recordings of each group interview session	▪ Preliminary list of fundamental calculus concepts ▪ Preliminary version of understanding frameworks ▪ Drafts of 19 tasks
Cycle 2: Analyzing and modifying tasks based on faculty testing and student work. (Implicit revisions of concept list and frameworks)	▪ Revision of tasks ▪ Revisions of concept list and frameworks	▪ Student work for 11 selected tasks ▪ Interview notes from each interdisciplinary group ▪ Modified tasks ▪ Audio recordings of each faculty group interview session	▪ Revised list of fundamental calculus concepts ▪ Revised version of understanding frameworks ▪ Revisions of 11 tasks

[2] The "faculty lens" is comprised of any pre-existing beliefs and/or knowledge about calculus, any previous experience with the task themselves or with similar tasks, and any work done to complete the prototype tasks.

Data Collection	Goals	Data Collected	Products
Cycle 3: Clarifying Distinctions, Evaluating Tasks	▪ Clarify and make explicit the distinctions between teachers and users ▪ Revision of tasks	▪ Student work for 12 selected tasks ▪ Interview notes from each interdisciplinary group ▪ Modified tasks ▪ Audio recordings of each faculty group interview session	▪ Final list of fundamental calculus concepts ▪ Final version of understanding frameworks ▪ Final versions of 12 tasks

Figure 2. Description of Cycles, Data Collected, and Products.

Results and Discussion

Before presenting the results and discussion for the interviews, one comment must be made regarding the chosen client discipline courses. As mentioned above, the client discipline courses were selected because Calculus I and/or Calculus II was listed as a pre- or co-requisite. Surprisingly, the instructors of several of these courses would not consider them to be calculus-intensive or even calculus-based courses.

- *Astronautical Engineering*: One faculty member declined participation because she felt the astronautical engineering courses of interest to me did not have "sufficient material to be of use" to this study because "at no point in the semester are students actually required to perform calculus operations, [although] we talk about derivatives."
- *Physics*: The physics courses are ramping up the incorporation of calculus. "In 10 years, we'll be able to say it's a calculus-intensive course. But certainly there's more calculus in it now than there was last year and more last year than the year before."
- *Operations Research*: The operations research faculty participants admitted that calculus is a pre-requisite "for mathematical maturity more than just the actual calculus" and because "the way [the course] is taught, you can do it without calculus."

These comments and others influence how many of the participants view calculus and the calculus understanding required for their particular courses.

The results and analyses discussed here are not in their raw form. All data interpretations and follow-on analyses were reviewed by the faculty participants to ensure completeness and accuracy. The quality, usefulness, and effectiveness of the prototype tasks (see the appendix for examples of the tasks written) were tested through administration to single-variable calculus students and analysis by the interdisciplinary groups of faculty participants. A complete cycle of data collection, data analysis, and interpretation verification occurred for a given session prior to conducting subsequent sessions.

Cycle 1: Describing Calculus and its Fundamental Concepts, Developing Draft Tasks

The first round of interviews began with a very general discussion of calculus and understanding and then progressed to the very specific. Each intradisciplinary interview (with 2-4 faculty participants per group) culminated in the faculty participants developing tasks to elicit the calculus understanding discussed in the sessions. For a detailed description of Cycle 1, see Ferguson and Lesh (2011).

One goal of design research is to put people with different perspectives into situations that require them to express not only how they think about a concept, but to express it in such a way that requires them to test and revise their way of thinking (Lesh, 2002). The iterative design of this study allowed this testing and revision to occur; however, the testing and revising did not occur until Cycles 2 and 3. To establish a baseline, Table 1 summarizes the indicators of calculus understanding the faculty participants exhibited and/or articulated during Cycle 1. Note: Red text describes procedural aspects and blue text describes conceptual aspects. An "X" means the faculty participant group (teachers or users) exhibited and/or articulated this aspect. The bottom three rows summarize the faculty participants' view of calculus, their list of the fundamental calculus concepts (given the context of their respective courses), and their chosen method to elicit calculus understanding from students (as indicated by the tasks written during this cycle).

Table 1
Cycle 1 Indicators of Different Aspects of Calculus Understanding (Procedural and Conceptual), by Faculty Participant Grouping

Indicators of different aspects of understanding (procedural and conceptual) by faculty participant grouping	**Teachers of Calculus**	**Users of Calculus**
Procedural Fluency (i.e., do a procedure; use a calculus tool)	X	X
Procedural Explanation (i.e., use words, not numbers or symbols, to do a procedure)	X	
Pattern-matching (i.e., recognize situations that differ in number or variable name as the same)	X	X
Language and Notation Usage (i.e., use of "critical" words to signal expertise)	Abstract meaning stressed	Physical meaning stressed
Tool Selection (i.e., analyze/assess situation to decide which prefabricated tool is appropriate)	X	X
Procedural Application (i.e., justify/defend tool selection)	X	
Interpretive and Predictive Power (i.e., interpret the solution in context) (i.e., analyze/predict the effect of changing the variable's value)		X X
View of Calculus (understanding) (a) computation-focused (b) conception-focused	(a) & (b) Settled on *calculus is a tool*	(a) & (b) Settled on *calculus is a tool*
Identification of important ideas and concepts	Function Derivative Integral Limit	Function Derivative Integral Limit (engineering only)
Elicitation Method and *Context* (i.e., chosen method and context to elicit calculus understanding)	Computations coupled with explanations in abstract or theoretical contexts (*contrived*)	Computations in physical contexts (*contrived*)

According to the teachers of calculus, understanding is about *how* and *why*. At the end of each calculus course, teachers of calculus want students to walk away with a toolbox full of tools, or procedures, the students know *how* to use, as well as *why* they should use them. Teachers of calculus want students to develop *procedural fluency* (i.e., to be able to carry out pre-fabricated procedures flexibly, accurately, efficiently, and appropriately). They also want students to develop *procedural application* (i.e., a student is able to discuss the pros and cons of a procedure, what is needed to apply one procedure versus another, and what procedure is appropriate). They believe "learning skills leads to building concepts." As such, the teachers of calculus elicit calculus understanding via computational questions which are in tandem with explanations.

According to the users of calculus, understanding is about *which* and *what is happening*. Within the client courses, the users of calculus want students to have more than just a toolbox full of tools, or procedures: "It's not so much that [the students] understand how to turn the crank and spit out an answer. Really mastering [calculus] relies on understanding what that integral or what that derivative actually means in the physical world." Users of calculus want students to develop *relational understanding* (Skemp, 1976/1978/2006; i.e., recognition of the concept being dealt with and relates it to an applicable procedure). They also want students to develop *predictive power* (i.e., a student is able to step out of the mathematics and recognize that in the mathematics, there is a prediction or truth about what is happening in the physical world). Students need to assess physical situations and select the calculus tool (i.e., a pre-fabricated concept and associated procedure) that will enable them to make sensible predictions about the situation. For example, "what does the variable in this equation, that the student just constructed, mean? If it doubles, what happens to the real world? Does the student suddenly get a space ship that flies faster than the speed of light? If so, something must be wrong with the mathematics." The users of calculus chose to elicit calculus understanding via computational questions.

According to both teachers and users of calculus, understanding is assessing the given (contrived) situation and intelligently selecting an existing tool of the expected concept and applying the associated procedures correctly to get a reasonable answer and/or prediction. Students must also be able to justify and defend their choice of tool(s).

Thinking of understanding can be likened to cooking. At the end of Cycle 1, the articulations of what the teachers and users of calculus think of understanding can be likened to the skills of a beginning cook. A beginning cook will go to the cupboard and assess the ingredient situation – say, they find tomatoes, mozzarella, and oregano. To this cook, this means an Italian meal and so he/she locates an Italian cookbook and selects a recipe that uses only the ingredients in the cupboard. The recipe is then followed, step-by-step, and a meal is produced. However, the teachers and users of calculus do look at understanding a bit differently, similarly to the different skills beginning cooks possess. The teachers' view of understanding would be similar to the way a beginning cook can talk his/her dinner guests through how the meal was made and maybe even be able to discuss why the recipe chosen was the best based on the given ingredients. On the other hand, the users are more interested in whether their students have an understanding that allows them to use calculus similarly to how some beginner cooks can double or half the recipe and produce a good meal. Basically, both of the beginner cooks described here have mostly procedural understandings of cooking; with a little dabble of knowing why the recipe (i.e., procedure) was appropriate and how to make minor changes (e.g., using individual basil, marjoram, oregano, and sage when the recipe calls for Italian seasoning) to the recipe if needed.

Cycle 2: Analyzing and Modifying Tasks based on Faculty Testing and Student Work

For the second round of interviews, the faculty participants were put into interdisciplinary groups, with 2-5 faculty participants per group. Each interview session began with a general discussion of whether a task must involve calculus computations to elicit calculus understanding. Following this discussion, the faculty participants evaluated the eleven tasks. Tasks were evaluated and analyzed first through the faculty lens and then through the medium of student work. Task modification completed the interview session. Implicit in the task modification discussions lay opportunities for revisions of the faculty participants model(s) of calculus and the concepts that comprise the field, as well as revisions of what it means to understand calculus.

The months between Cycle 1 and Cycle 2 allowed not only for the researcher to review, analyze, and interpret the data from Cycle 1, but also for the faculty participants to really think about calculus understanding from the perspective of this study. Discussing the student participants' work on the tasks revealed the revisions the faculty participants' thinking had undergone. The transition in their thinking is mapped in Table 2. Note: Gray text describes indicators specific to any previous cycle.

Table 2

The Addition of Cycle 2's Indicators of Different Aspects of Calculus Understanding (Procedural and Conceptual), by Faculty Participant Grouping

Indicators of different aspects of understanding (procedural and conceptual) by faculty participant grouping	**Teachers of Calculus**		**Users of Calculus**	
	Cycle 1	Cycle 2	Cycle 1	Cycle 2
Procedural Fluency (i.e., do a procedure; use a calculus tool)	X		X	X
Procedural Explanation (i.e., use words, not numbers or symbols, to do a procedure)	X	X		
Pattern-matching (i.e., recognize situations that differ in number or variable name as the same)	X		X	
Language and Notation Usage (i.e., use of "critical" words to signal expertise) (i.e., interpretation on meaning based on contextual clues)	Abstract meaning stressed		Physical meaning stressed	
Tool Selection (i.e., analyze/assess situation to decide which prefabricated tool is appropriate)	X		X	X
Procedural Application (i.e., justify/defend tool selection)	X			
Interpretive and Predictive Power (i.e., interpret the solution in context) (i.e., analyze/predict the effect of changing the variable's value)			X X	X
Relational Connection (i.e., increase and deepen meaning of concepts by developing different ideas and viewpoints of concepts)		X		X
View of Calculus (understanding) (a) computation-focused (b) conception-focused	(a) & (b) Settled on *calculus is a tool*	*Calculus is a tool* used to 1. Describe how to solve problem 2. Solve problems	(a) & (b) Settled on *calculus is a tool*	*Calculus is a tool* used to 1. Explain how a physical situation works 2. Make predictions 3. Solve problems
Identification of important ideas and concepts	Function Derivative Integral Limit	Function Derivative Integral Relationship between derivative and integral	Function Derivative Integral Limit	Function Derivative Integral Relationship between derivative and integral
Elicitation Method and *Context* (i.e., chosen method and context to elicit calculus understanding)	Computations coupled with explanations in abstract or theoretical contexts (*contrived*)	*Explanations*, with occasional computations, computations are not required (*contrived*)	Computations in physical contexts (*contrived*)	*Applications:* computations in some physical context (*contrived*)

Comments made by the teachers of calculus during Cycle 1 suggested a focus on developing the students' knowledge of how to use the different calculus tools and their abilities to explain what they are doing. The teachers' comments throughout Cycle 2's discussions and evaluations of student work on the faculty-authored tasks, as well as the researcher-authored and selected tasks, evidenced a shift in this focus. While computations are a necessary part of the calculus curriculum, the teachers expressed a preference to focus on the explanation abilities of the students; and as such, choose to assess understanding with explanations and descriptions (e.g., writing assignments and setting up of a model).

To the teachers' articulated focus on explanations, the users of calculus responded with this question: While students have to understand the concept to be able to write about it, does that also mean they can apply the concept? Application has been the explicit focus of the users of calculus throughout both Cycle 1 and 2. User comments during Cycle 2's discussions revealed an emphasis on the applicative nature of the calculus tools the users need students to work with in order to succeed in their respective discipline courses. Application of calculus tools in the client discipline courses requires the students to recognize 1) that calculus is applicable to the given situation and 2) which calculus tool is appropriate and when to use it. Therefore, the users of calculus choose to assess understanding with applications. This difference in focus could explain why there is a mismatch between the actual preparation the students receive in the calculus courses and the preparation the users of calculus expect.

Elevating the relationship between derivatives and integrals to a fundamental concept illuminated a big shift in the faculty participants' thinking about calculus and how they elicit calculus understanding: from developing concepts in isolation to focusing on the relational connection(s) between the concepts. Students increase and deepen their understanding of the individual concepts by developing different ideas and viewpoints of the concepts. These different ideas and viewpoints come from exploring the relationships between the concepts and evaluating how the concept(s) interact with the given situations and contexts. The teachers' and users' use of contrived situations and contexts limits the exploration and evaluation opportunities a student is afforded. Both the teachers and users recognized the potential of using more realistic situations and contexts to develop and assess calculus understanding when they evaluated the student work on the MEA-type tasks.

The revisions of the faculty participants' ways of thinking about calculus understanding show an emerging progression from the beginning cook (described in the **Cycle 1** section) that cannot modify a recipe to accommodate on-hand ingredients to a slightly more advanced cook that can modify a recipe because he/she knows how different ingredients interact and react to each other to create a delicious meal. This second cook still depends on a recipe, but is able to substitute similar ingredients (e.g., substituting cottage cheese for ricotta cheese, to reduce the grittiness of a lasagna) and still produce a delicious meal.

Cycle 3: Clarifying Distinctions, Evaluating Tasks

Cycle 3 proceeded much like Cycle 2 with Calculus I and Calculus II student work sessions and faculty interview sessions. The set of 14 tasks administered during the student work sessions consisted of refined and modified versions of Cycle 2's tasks and one new task written by a faculty participant. Faculty participant interviews (the faculty participants were regrouped into different interdisciplinary groups, with 2-4 faculty participants per group) followed during which the distinctions between how the faculty participants thought about calculus understanding and how to best elicit it from students were clarified and articulated. Interwoven throughout these discussions were evaluations and analyses of the tasks, again evaluated and analyzed through the faculty lens and the medium of student work.

The following statements articulate and distinguish the "parallel" thinking demonstrated during this study by the teachers and users when it came to objectives for understanding, application, awareness, extraction, guidance, and representation.

1. *Understanding*: The distinction between how the teachers and users view understanding calculus (beyond computations) is best articulated by the level of understanding students are expected to demonstrate upon leaving their calculus course(s) or entering their client course(s). Using Bloom's taxonomy (Anderson & Krathwohl, 2001), these levels are:
   - Understanding: Can the student explain ideas and concepts? (teacher preference)
   - Applying: Can the student use the information in a new way? (user preference)

   Because the teachers of calculus are focused on developing explanatory abilities in their students, they have little to no time for applications. This reality does not meet the users' expectation (or desires) for students to be steeped in applications.
2. *Application*: For the teachers, understanding a concept and applying a concept are different and exist in hierarchy. Application without understanding is repetition of a teacher- or textbook-demonstrated procedure. Application with understanding is being able to "undress" the given situation, recognize the underlying concept(s), and select the appropriate tool for solving the problem. The teachers believe computations are required to elicit application, while they are not needed to elicit understanding. For the users, application is the ability to apply a concept with understanding (i.e., recognizing the concept within a new situation, knowing what procedures then apply, and proficiently solving for the answer). Computational ability and versatility must combine with understanding to get the ability to solve novel problems.
3. *Awareness*: Several Cycle 2 comments opened a discussion of the necessity for students to *think* they are doing calculus when they are doing calculus. While all the faculty participants agreed it would be beneficial for students to recognize what was causing their difficulties when solving a problem (e.g., deficit in algebra not calculus when trying to optimize), they differed on whether it was important to label the tools that allowed the problem to be solved. For some, the ability to label the tools is completely unnecessary; while for others, the ability to label is synonymous with selecting the appropriate tool.
4. *Extraction*: The users expect the students to be steeped in applications – applications with understanding – when they leave the pre-requisite calculus courses. As a result, computations (i.e., numerical or graphical solutions) are mandatory. Meanwhile, the teachers of calculus focus on developing explanatory abilities in their students. While explanations do a good job of eliciting understanding, simultaneously eliciting understanding and mechanics is the agreed upon "best" method for elicitation.
5. *Guidance*: The amount of "leading" a task must do to either guide a student in the desired direction or determine where a student is having difficulties depends on the timing and purpose of the task. The teachers felt the students need more guidance because the purpose of their instruction is to help students develop the procedures and concepts; it is the users' job to help students develop the flexibility to apply the procedures and concepts. The users did not disagree, but stressed the use of guidance depended on the type of assessment: formative meant leading the student and summative meant "throwing the students out of the nest and seeing if they could fly."
6. *Representation, specifically tables versus graphs*: While the faculty participants viewed each representation as merely a different way to look at a situation, they felt working with a table requires more assumptions and thus requires a deeper understanding than working with a graph does. The teachers emphasized that because most students do not realize the assumptions necessary to work with tables, they (the teachers) prefer graphs. Also impacting this choice is the abstract nature of their favored situations (i.e., situations lending themselves easily to continuous, smooth graphs). The users prefer tables because

most real-world data are collected, stored, and displayed using tables; real life does not typically have very nice and neat functions associated with it; and making the assumptions explicit is essential to the students' understanding of the situation.

As the months between Cycles 1 and 2 did, the time between Cycles 2 and 3 allowed for data analysis by the researcher and in-depth thinking by the faculty participants. Further revisions in how the faculty participants think about calculus understanding were revealed throughout Cycle 3's review, clarification, and articulation of this study's conclusions and throughout the review of student work. The transition in their thinking is mapped in Table 3. Note: Purple text describes application aspects (i.e., the melding of procedural knowledge and conceptual understanding).

Table 3
The Addition of Cycle 3's Indicators of Different Aspects of Calculus Understanding (Procedural and Conceptual), by Faculty Participant Grouping

Indicators of different aspects of understanding (procedural, conceptual, and application) by faculty participant grouping	**Teachers of Calculus**			**Users of Calculus**		
	Cycle 1	Cycle 2	Cycle 3	Cycle 1	Cycle 2	Cycle 3
Procedural Fluency (i.e., do a procedure; use a calculus tool)	X		X	X	X	X
Procedural Explanation (i.e., use words, not numbers or symbols, to do a procedure)	X	X	X			X
Application Ability (i.e., transfer knowledge to different situations with structural similarities)			X			X
Pattern-matching (i.e., recognize situations that differ in number or variable name as the same)	X			X		
Tool Selection (i.e., analyze/assess situation to decide which prefabricated tool is appropriate)	X			X	X	
Procedural Application (i.e., justify/defend tool selection)	X					
Communicating Calculus (i.e., communicate procedures, reasonings, solutions, generalizations of calculus tool using appropriate technical and contextual language)			X			X
Language and Notation Usage (i.e., use of "critical" words to signal expertise) (i.e., interpretation on meaning based on contextual clues)		Abstract meaning stressed			Physical meaning stressed	
Interpretive and Predictive Power (i.e., interpret the solution in context) (i.e., analyze/predict the effect of changing the variable's value)				X X	X	
Relational Connection (i.e., increase and deepen meaning of concepts by developing different ideas and viewpoints of concepts)		X			X	

Indicators of different aspects of understanding (procedural, conceptual, and application) by faculty participant grouping	**Teachers of Calculus**			**Users of Calculus**		
	Cycle 1	Cycle 2	Cycle 3	Cycle 1	Cycle 2	Cycle 3
View of Calculus (understanding) (a) computation-focused (b) conception-focused	(a) & (b) Settled on *calculus is a tool*		*Calculus is a tool* used to 1. Describe how to solve problem 2. Solve problems	(a) & (b) Settled on *calculus is a tool*		*Calculus is a tool* used to 1. Explain how a physical situation works 2. Make predictions 3. Solve problems
Identification of important ideas and concepts	Function Derivative Integral Limit	Function Derivative Integral Relationship between derivative and integral		Function Derivative Integral Limit	Function Derivative Integral Relationship between derivative and integral	
Elicitation Method and *Context* (i.e., chosen method and context to elicit calculus understanding)	Computations coupled with explanations in abstract or theoretical contexts (*contrived*)	*Explanations*, with occasional computations, computations are not required (*contrived*)	Simultaneous explanation and solution (MEA-type tasks)	Computations in physical contexts (*contrived*)	*Applications:* computations in some physical context (*contrived*)	Simultaneous explanation and solution (MEA-type tasks)

As can be seen in Table 3, a trend in the faculty participants' thinking is the progressive articulation of the need for and stress put on explanations. Five of the eight original faculty-authored tasks required *only* a numerical or graphical solution. At the end of Cycle 3, three of the eleven polished non-MEA-type tasks required only a numerical or graphical solution. The ability and skill required to apply calculus at the level the users of calculus require obligates the use of tasks that ask the students to do both the mechanics and explanations simultaneously. According to the faculty participants, eliciting students' understanding in this manner is the best of all the worlds because it 1) exposes the area(s) where the student is struggling, 2) elicits really good information about what the student understands, and 3) asks the student to do the calculus.

As the discussions progressed, the faculty participants exhibited a greater preference for conceptual aspects of understanding rather than procedural aspects in their articulations of what it means to understand calculus and what it means to apply calculus. Application is not just a melding of understanding and procedural knowledge, it is taking that understanding and internalizing it and using it in a new way (i.e., in situations with structural similarities). This unified definition shows that while the faculty participants talked understanding, they had one foot firmly planted in procedural knowledge. The students MUST be able to correctly execute a procedure. One user of calculus put it quite bluntly:

> I don't really care whether [students] understand [calculus]. Understanding why a derivative works doesn't do me any good; they've got to start applying it because calculus doesn't help me unless it's allowing them to analyze the real world and get a sense of what's happening. … It's that level of application and what it means in the real world – in scenarios – that I care about. *(Engineering faculty)*

Not only did this engineering faculty emphasize the need for students to be able to correctly execute a procedure, she implied the need for a level of understanding that combines all the aspects of understanding that similarly allow a more advanced cook to assess the ingredient situation and compose a recipe instead of having to choose one. This person is always going to be a better cook because he/she can take whatever is in the refrigerator, what is fresh and in season, and create an excellent meal.

This more advanced level of cook has the type of flexible, durable, and applicable understanding the faculty participants want from calculus students. They want a student that can use calculus to describe a realistic situation, one that might not fit any existing descriptions, textbook examples, or library entries. When a student arrives at a result, they will know if and how it applies to the given situation and whether it applies to any other situation(s). This occurs because they know what assumptions they made, what error or uncertainty might be involved. Because the calculus description was created by the individual, he/she owns it. This personal ownership enables him/her to not only build and deepen their understanding, but to apply calculus in novel and unfamiliar situations of any kind. This type of understanding is flexible, durable, and transferrable. One faculty participant summed it up this way: "Ultimately we want [students] to not just be able to do problems similar to what they've seen before, we want them to take what they've known and do something new."

Implications and Conclusion

Teachers of calculus push students to develop a deep understanding of the concepts such that they can explain it to another student well enough to make that student understand; whereas the users expect the students to walk into their classrooms with a deep understanding of the concepts such that they can recognize a concept within the given situation or context, select a calculus tool that will efficiently get them an answer, compute the answer, and make an accurate determination or prediction. As this study revealed, the end goal of the calculus courses and the beginning goal of the client disciple courses do not align. This misalignment has caused and continues to cause many students, instructors, and researchers much frustration. This misalignment has caused and continues to cause many students, instructors, and researchers much frustration (Ferrini-Mundy & Graham, 1991; Klingbeil, et al., 2006). This misalignment also carries with it many implications that have already begun to occur: implications such as client disciplines creating mathematics courses that are taught in-house by client-discipline faculty, which will remove students from mathematics classrooms; thoughts that mathematics classrooms should teach only mechanics and leave the bridging of the procedures and context (i.e. the applications) to the client discipline courses; or questions about the purpose served by the Calculus I and Calculus II courses. If one of the main purposes of the calculus courses is to prepare students for success in the calculus-based client-discipline courses, then as the faculty participants came to realize, there are several courses of action that the faculty of all the disciplines can pursue.

Decades ago when researchers began investigating and characterizing student understanding of many calculus concepts, they found "exceptionally primitive" understandings and a lack of intuition about the concepts and procedures of calculus (Ferrini-Mundy & Graham, 1991, p. 631). Not much has changed in the intervening time. Mathematics and client discipline teachers are still expressing concerns about the apparent lack of understanding of calculus, especially when students are asked to use it in unfamiliar situations (Ganter & Barker, 2004; Hughes Hallett, 2000). What has changed in the intervening years is a development of frameworks to better characterize student understanding of the fundamental calculus concepts. What has not changed is the fundamental structure of the underlying curriculum (Thompson, Byerley, & Hatfield, in press). "It seems that, while the reform of calculus has had an impact on calculus rhetoric, it has not had an impact on what is expected that students learn"(Thompson, et al., in press, p. 3). One possible course of action to align the expectations is to consider radically different curriculum. Examples are Thompson et al.'s (in press)conceptual approach to calculus using the Fundamental Theorem of Calculus as the cognitive root and Samuel's (2012) approach to differential calculus using local linearity as the cognitive root.

Whether new curriculum is explored or not, calculus instructors need to break away from "traditional" application problems. These problems typically refer to the word problems found at the end of textbook chapters that present students with a real-world context and require that students ascertain which mathematics tool (and associated procedure) they should apply. Some research suggests that although students may become proficient at executing procedures efficiently through traditional instruction, they often have a poor conceptual understanding of the fundamental concepts (Eisenberg & Dreyfus, 1991; Thomas & Hong, 1996; Thompson, 1994). Research much like this and the experiences of the teachers of calculus (anecdotal evidence) has prompted them to seek feedback from the client disciplines and to modify their curriculum and instruction based on that feedback. The teachers of calculus "feel like [they] have an obligation, as a Math Department, to teach; but [they] also have an obligation to send [the students] off to their engineering classes and physics classes able to do those kinds of things." However, the majority of the calculus instructors are still using traditional application problems and problem solving techniques and methods to prepare the students for the client discipline courses. These choices are contributing to the gap between what experiences are provided to students in the calculus courses and the experiences the users of calculus expect the students to have when entering their respective classrooms.

This research is also reason to not revert back to teaching only procedures (i.e., traditional instruction) as one teacher of calculus felt some of the client discipline participants had implied. The calculus reform efforts have been and continue to motivate calculus instruction to bridging the gap between knowing how to use the tool versus knowing how and why the tool works and recognizing when to apply it. One way to bridge the gap is to supplement the traditional application problems with tasks, such as the prototype tasks produced by this study. Using modeling problems, problems that require students to make mathematical descriptions of meaningful situations, can "dare teach more than any teacher has thought possible" (Lesh & Doerr, 2003, p. 5).

As the teachers and users at the institution where this study was conducted have learned to do and have experienced as beneficial, a consistent and reflective dialog should be opened between the mathematics and client disciplines to ensure instructors are preparing students for their further academic endeavors in whichever discipline they may pursue. Part of this dialog should focus on constructing mutual definitions f what it means to understand calculus and what it means to apply calculus. Other portions of this dialog can focus on the concepts and topics within the calculus courses to make explicit what is and is not part of the curriculum. Then further discussions can center on how these concepts and topics are taught and used in the various disciplines. Most importantly this dialog needs to be reflective and action-spurring.

As stated before, calculus is an important tool for building mathematical models of the world around us and is thus used in a variety of disciplines, such as physics and engineering. These disciplines rely on calculus courses to provide the mathematical foundation needed for success in their discipline courses. This study hopes to offer a collective vision to focus the content of beginning calculus courses on the meeting the needs of client disciplines. However, it is ultimately the mathematicians that have the responsibility to create courses and curricula that embrace the spirit of this vision while maintaining the intellectual integrity of mathematics. By explicitly knowing what and how students should be prepared for client courses, teachers and curriculum developers of both calculus and client disciplines can work together to prepare students for academic success in any discipline.

References

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives, Complete edition*. New York, NY: Longman.

Bressoud, D. M. (1992). Why do we teach calculus? *The American Mathematical Monthly, 99*(7), 615-617.
Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141-178.
Carlson, M., Oehrtman, M., & Engelke, N. (2010). The precalculus concept assessment: A tool of assessing students' reasoning abilities and understandings. *Cognition and Instruction, 28*(2), 113-145.
Collins, A. M. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22). Berlin; New York: Springer-Verlag.
Eisenberg, T., & Dreyfus, T. (1991). On the reluctance to visualize in mathemtics. In W. Zimmermann & S. Cunningham (Eds.), *Visualization in teaching and learning mathematics* (Vol. MAA Notes, No 19, pp. 25-38). Washington D.C.: The Mathematical Association of America.
Ferguson, L., & Lesh, R. A. (2011). Concepts fundamental to an applicable understanding of calculus. In S. Brown, S. Larsen, K. Marrongelle & M. Oehrtman (Eds.), *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education* (Vol. 1, pp. 50-67). Portland, OR.
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*(7), 627-635.
Ferrini-Mundy, J., & Gücler, B. (2009). Discipline-based efforts to enhance undergraduate STEM education. *New Directions for Teaching and Learning, 117*, 55-67.
Ganter, S. L., & Barker, W. (2004). A collective vision: Voices of the partner disciplines. In MAA (Ed.), The Curriculum Foundations Project (pp. 1-13). Washington, D.C.: Mathematical Association of America's Committee on the Undergraduate Program in Mathematics.
Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., . . . Human, P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.
Hughes Hallett, D. (2000). *Calculus at the start of the new millennium.* Paper presented at the Proceedings of the International Conference of Technology in Mathematics Education, Beirut, Lebanon.
Kasten, M. (1988). The role of calculus in college mathematics. *ERIC/SMEAC Mathematics Education Digest No. 1.* Retrieved from (ERIC Publications Service No. ED321970).
Klingbeil, N. W., Mercer, R. E., Rattan, K. S., Raymer, M. L., & Reynolds, D. B. (2006). *The WSU model for engineering mathematics education: Student performance, perception and retention in year one.* Paper presented at the 2006 ASEE Illinois-Indiana and North Central Joint Section Conference, Fort Wayne, IN.
Lesh, R. A., & Doerr, H. M. (2003). Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. A. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp. 3-33). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Lesh, R. A., & Kelly, A. E. (2000). Multi-tiered teaching experiments. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 197-230). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Samuels, J. (2012). *The effectiveness of local linearity as a cognitive root for the derivative in a redesigned first-semester calculus course.* Paper presented at the 15th Annual Conference on Research in Undergraduate Mathematics Education, Portland, OR.
Thomas, M. O. J., & Hong, Y. Y. (1996). *The Riemann integral in calculus: Students' processes and concepts.* Paper presented at the 19th Mathematics Education Research Group of Australasia Conference, Melbourne, Australia.
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., Byerley, C., & Hatfield, N. (in press). A conceptual approach to calculus made possible by technology. *Computers in the Schools*.
Yoon, C., Dreyfus, T., & Thomas, M. O. J. (2010). How high is the tramping track? Mathematising and applying in a calculus model-eliciting activity. *Mathematics Education Research Journal, 22*(1), 141-157.
Zerr, R. J. (2010). Promoting students' ability to think conceptually in calculus. *PRIMUS, 20*(1), 1-20.

Appendix

The Ball Task (modified from a composition contextual problem discussed in Carlson, Oehrtman, & Engelke, 2010) and the Hiking Task (modified from the Tramping Problem in Yoon, Dreyfus, & Thomas, 2010) are offered here as examples of the tasks used in this study. Both tasks are in the post-Cycle 3 form.

The Ball Task

Consider the following problem: A ball is thrown into a lake, creating a circular ripple that travels outward at a speed of 5 cm per second. How much time will it take for the area of the circular ripple to exceed 50,000 square centimeters?

Write up a solution to the above problem as if you were writing a textbook example (i.e., explain how any function, formula, and/or computation will be used in the next step(s)).

For example: *Find the second derivative and interpret its sign for □□=□3.*

Solution: If □□=□3, then □'□=3□2, □''□=□□□3□2=3□□□□2=3·2□=6□. This is positive for □>0 and negative for □<0, which means □□=□3 is concave up for □>0 and concave down for □<0.

The Hiking Task

To celebrate their 40th wedding anniversary, Helen and Brendan O'Neill are planning a hike with their children and grandchildren. They are considering a nearby 5-kilometer hike. The local park provided a graph of the trail's grade at every point, but the O'Neills want to make sure it is suitable for them. Helen wants to know if there is a summit where they can have lunch and enjoy the view, while Brendan wants to know where the hiking gets difficult.

The O'Neills need your help!

Design a method that the O'Neills can use to sketch a distance-height graph of the original trail. You can assume the trail begins at sea level.

Write a letter to the O'Neills explaining your method, and use your method to describe what the hiking trail will be like. In particular, you must clearly show any summits and valleys on the trail, uphill and downhill portions of the trail, and the parts of the trail where the slopes are steepest and easiest.

Most importantly, your method needs to work not only for this hiking trail, but also for any other hiking trail the O'Neills might consider.

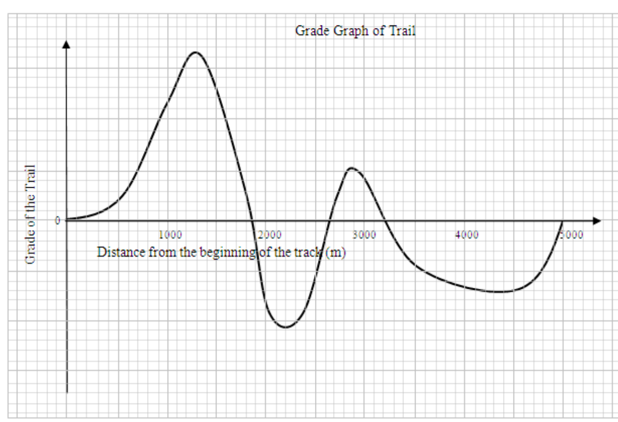


Figure 3. Two prototype tasks: the Ball and the Hiking Tasks.

THE IMPLIED OBSERVER OF A MATHEMATICS LECTURE

Tim Fukawa-Connelly[1], Aaron Weinberg[2], Emilie Wieser[2], Sarah Berube[1] and Kyle Gray[1]
1. Department of Mathematics and Statistics, The University of New Hampshire
2. Department of Mathematics, Ithaca College

There is a need to explain the relationship between teaching (classroom activities) and the resulting student learning, especially in advanced mathematics classes. This study represents a first attempt to describe the opportunity to learn present in an abstract algebra lecture, as an exemplar of advanced mathematics. Based on Weinberg and Wiesner's (2010) work on the implied reader of a mathematics textbook, we describe the implied observer of a lecture as a bundle of codes, competencies and behaviors that are needed to make a meaningful interpretation of the lecture. We also use the framework to analyze an abstract algebra class and describe needed codes, competencies and behaviors of a student of that class. Finally, we close the paper by discussing both theoretical and methodological questions that remain and how those questions give rise to disagreements about how to interpret a component of the lecture by expert observers.

Key words: implied observer, codes, competencies, behaviors, abstract algebra

Background and Motivation.

The dominant format of instruction in college classrooms is the lecture (Armbuster, 2000). In advanced undergraduate mathematics courses, the lecture is likely to center on the presentation of definitions, theorems, proofs, and examples. In order to determine how students learn in such courses, it is important to understand how various aspects of the lecture relate to what students can and do "take away."

There have been a variety of studies focused on activity in the undergraduate classroom and its implications for learning. Kiewra (1991) and Titsworth (2004) investigated how various instructor practices (including the use of organizational cues, preparation of note-taking guides, and levels of immediacy) affect students' notes. Fukawa-Connelly (in press) described pedagogical strategies used by the instructor of an abstract algebra class; this description focused on how the instructor modeled expert mathematical behavior in the process of constructing a proof. Mills (2012) examined the relationship between instructors' pedagogical intentions and practice in proof-based undergraduate mathematics courses.

Although previous studies offer a variety of perspectives on what happens in college classrooms, they do not provide a complete set of tools for analyzing the moment-to-moment experience of a student in a lecture. In this paper, we propose a framework for doing so. The framework has two parts: a classification of the lecture content in terms of mode of presentation and purpose; and a description of the implied observer of the lecture (that is, the collection of demands that lecture places on the student in order to fully participate in the lecture). In establishing this framework, we draw on Shein's (2012) work on mathematics teachers' use of gesture, the framework for proof comprehension developed by Mejia-Ramos et al. (2010), and Weinberg and Wiesner's (2011) description of the implied reader of a mathematics textbook.

Introductory Example.

This paper focuses on a three-minute excerpt (transcribed below) from an undergraduate abstract algebra course. The course was taught at a mid-size public research university and is typically taken by junior and senior mathematics and mathematics education majors. The lecturer had considerable previous experience teaching undergraduate abstract algebra. Lectures followed the traditional format of a mix of definitions, theorems, proofs, and examples, and the lectures did not include interaction between the lecturer and students. This excerpt comes from a lecture midway through the semester. It begins shortly after the start of the lecture and, typical of the instructor's lectures, is a review of the final few moments of the previous lecture.

In examining this transcript, we ask the following questions: What are the essential components of the lecture? What opportunities to learn does the lecture provide, and how are these opportunities constructed by the lecture components? In order to take advantage of these opportunities, what is demanded of students? We will address these questions by examining the components of the *implied observer* of a lecture and how these components affect students' opportunity to learn.

Lecture Content (organized by component)

Time-stamp	**Spoken**	**Written**	**Gesture**
12:17	You recall—so I gave a proof last time.		
12:20	It was *by determinants.*	`Pf: By determinants.`	
12:27	And I'm just going to let this *stand*.		Holds hand under the phrase "By determinants".
12:30	You'll recall the general idea was this.		
12:37	That we took at a fact from linear algebra, that if you…		
12:37	Have a matrix		Holds up hands with fingers vertical & thumbs pointed in
12:39	And let's recall what a determinant is,		Moves left hand in circle, right hand remained in prior position
12:41	right, for the sake of refreshing our memory on this.		
12:43	In the case of a 2x2 matrix, the determinant is given by a quite explicit formula. Well, it always is a quite explicit formula	[Writes a generic 2x2 matrix on right side of board]: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ [To left of matrix]: `det`	
12:55	but here it's somewhat simple because of the small size….		Points to matrix, then holds up hands with fingers vertical.
12:59	And, the observation that we wanted to make was that		

13:08	if you *change* the *sign* of a *row* or *column*	$\begin{bmatrix} -a & -b \\ c & d \end{bmatrix}$	
13:11	that… the *determinant* will also…	[On left of matrix]: `det` [To right of matrix]: =	
13:19	Don't do it that way… This is true…	[Erases matrix entries]	
13:23	If you *interchange* the… a pair of rows or a pair of columns, the determinant will change signs.	[Changes entries to]: $\begin{bmatrix} c & d \\ a & b \end{bmatrix}$	
13:32	And so, in this case, here, you see now that we get *c b* minus *a d*, which is the *negative of* ad minus bc.	[after the =]: ***`cb-ad = -(ad-bc)`***	
13:42	And there's the sign switch. Alright?		Moves hand back and forth in front of matrix
13:51	And so you'll recall, the idea was… to… think		Draws horizontal line underneath pair of matrices
13:53	—or to *define* what was mathematically called *action*, the action of… a… group on a… set.		Holds one hand out and cupped up, then rotates his arm in a semi-circle while flipping over his hand.
14:04	Or, in this case… we might want the… not be quite so specific		
14:08	but… it will be the action of a *Symmetric group*		Points toward left side of board
14:12	on the *set of matrices*… where, for example		
14:16	just to *borrow* again Cycle 2514 there		Points toward other side of the board.
14:22	That might act on a matrix M in the following way…	`(2514)M`	

Table 1: An annotated transcript of an abstract algebra class

The Lecture Components: The "Facts of the lecture."

We begin by establishing a general framework for describing the lecture content. This includes a characterization of the objective experience of the lecture—that is, the lecture components—in terms of the lecturer's modes of delivery (speech, writing, and gesture) and their relationship to each other in space and time. The framework also describes each lecture component in terms of its role in the lecture; that is, each lecture component may have a mathematical aspect or a communicational aspect.

There are three basic modes through which the instructor delivers a lecture: speech, writing, and gesture. (Note that the excerpt above already reflects this perspective.) *Speech* encompasses the words that the instructor says, emphasis and tone, and any pauses in speech. *Writing* includes all of the marks that the instructor makes on the board. Lastly, the lecturer may use *gesture* (that is, physical movement of his or her body) as part of the lecture content.

Shein (2012) has further classified gesture in terms its function. Pointing gestures are "used to indicate objects, locations, inscriptions, or students" (p. 185); pointing gestures focus observers' attention and help observers to follow the lecturer's words. For example, at 12:27, the lecturer holds his hand under the phrase "by determinants" to emphasize his speech and writing. Representational gestures are those "in which the handshape or motion trajectory of the hand or arm represented some object, action, concept or relation" (Shein, 2012, p. 186). In mathematics lectures, representational gestures may be used to portray a mathematical object or process. For example, at 13:53, the instructor uses a gesture to represent a group action. Lastly, Shein distinguishes writing gestures. We interpret writing as a gesture when the physical movement involved in writing conveys meaning (as opposed to the marks that result from writing.) For example, the lecturer may draw an arrow while creating a diagram; the movement of his hand while drawing the arrow helps to communicate the relationship between objects in the diagram.

In addition to describing the lecture in terms of its separate components, it's important to recognize how these components are coordinated. The experience of an individual piece of speech, writing, or gesture is partially determined by what comes before and after, chronologically, in the lecture as well as where it exists physically in the room. For example, at 13:51, the lecturer draws a horizontal line. This line shows a separation between the previous work establishing a property of the determinant and the use of this property as part of a larger proof. There are also likely to be multiple components occurring in a short period of time or possibly all at once. For example, the lecturer often writes at the same time that he speaks. The writing may duplicate what the lecturer is saying, may be a subset of what the lecturer is saying, or may contain different words or symbols than the lecturer's speech. Each of these situations may suggest a different type of coordination between the speech and writing. A match between speech and writing emphasizes the precise words that are being used. Writing that is a subset of speech suggests that the writing represents the essential ideas of the speech. Writing and speech that contain different words or symbols may provide different perspectives on the same idea. For example, at 13:32, the lecturer states "And so, in this case, here, you see now we get cb-ad, which is the *negative* of ad-bc" while writing out "bc-ad=-(ad-bc)." Here, the speech provides an interpretation of the significance of the symbolic written expression.

Each of the lecture components, along with their coordination in space and time, may have a communicational aspect or a mathematical aspect. Communicational aspects include organizational cues and immediacy. *Organizational cues* are lecture components that direct student attention (see, for example, Titsworth 2004). For example, the lecturer's use of the word "recall" at 12:27 directs students' attention to the content of the previous lecture. Emphasis in speech and pointing gestures also represent organizational cues. *Immediacy* describes lecturer behavior that reduces the social distance between the lecturer and their students (Mehrabian, 1971); the lecturer's use of gesture and tone and the use of the first-person plural (e.g. at 12:37, the speech "we took as a fact…") have an immediacy aspect.

Lecture components with a mathematical aspect include those related to mathematical facts, procedures, or processes. *Facts* include definitions, theorems, and examples. For example, at 13:53, the lecturer uses a gesture to represent a group action. *Procedures* are step-by-step methods that are presented or used in the lecture. *Processes* encompass problem-solving (e.g. creating and using a definition), mathematical communication or representation (e.g. using diagrams or formal symbols), and justification (e.g. use of formal proof structures). For example, at 12:43, the instructor selects a 2x2 matrix to illustrate an important property of determinants.

The implied observer.

Weinberg and Wiesner (2011) described the *implied reader* of a mathematics textbook as "the embodiment of the behaviors, codes, and competencies that are required of an empirical reader to respond to the text in a way that is both meaningful and accurate" (p. 52). This implied reader is distinct from the empirical reader—the person who actually reads the textbook—and from the intended reader—the characteristics of the reader that the authors have in mind. In particular, the implied reader is created by the text itself rather than by the people who write or read the book.

We define the *implied observer* of a mathematics lecture in the same way: the embodiment of the behaviors, codes, and competencies that are required of an empirical observer to respond to the lecture in a way that is both meaningful and accurate. The implied observer can be identified by expert observers as they reflectively examine the lecture. Although different expert readers may identify different attributes of the implied observer, the established conventions for mathematical communication and classroom lecturing allows the observers to share a common understanding of the lecture and, consequently, for these various attributes to contribute to a coherent description of the implied observer.

Codes. A code is the implied observer's method of ascribing meaning to particular lecture components. We classify codes into seven categories:

- *Formatting*, such as the layout of the board and the order in which components are presented. For example:
    - During the lecture, the instructor first writes a theorem and the beginning of the proof in the middle of the board. He then announces: "You'll recall the general idea was…" and then moved over to the far right side of the board. The implied observer interprets this physical shift as indicating that the next material is related to—but is not directly included in—the proof of the theorem.
- *Organizational Cues*, which are the words and actions that signal transition points in the lecture, such as switching from a summary of a previous lecture to introducing new material, or ending the proof of a theorem. For example:
    - After the instructor finishes discussing the effects of transposing rows of a matrix, he draws a horizontal line on the board and announced, "and so you'll recall, the idea was to think.…" Taken together, the implied observer interprets these as markers that the lecture is no longer discussing determinants but is transitioning to a related idea.
- *Symbols*, including commonly-used mathematical notation. For example:
    - At 12:41, the instructor writes a "generic" 2-by-2 matrix on the board using the letters *a*, *b*, *c*, and *d* as its entries. The implied observer interprets this configuration as a matrix and the letters as arbitrary elements of the ring.
    - At 13:32, the instructor writes "det" in front of the matrix and –(ad-bc) to the right of the matrix. The implied observer interprets this as taking and computing the determinant of the matrix. Furthermore, the implied observer interprets the first negative sign as indicating that the element in the parentheses is an additive inverse, while the second negative sign indicates the result of performing an action (specifically, *adding* the additive inverse).

- At 14:22, the instructor writes "(2514)M" on the board. The implied observer interprets the (2514) as a 4-cycle in a permutation group and M as an arbitrary matrix.

- *Diagrams*, which are usually charts and graphs, but may include other arrangements of symbols—including non-mathematical symbols (such as pictures).
- *Verbal cues*, such as pauses and forms of emphasis. In the example above, these typically occurred where the instructor either elongated the pronunciation of words, used vocal inflections (such as varying the pitch of his voice more than usual), or inserted pauses that weren't the result of filler words. For example:
  - At 12:20, the instructor emphasizes "by determinants." This indicates that this phrase is an important idea—in this case, it was the key idea or theme of the proof.
  - At 13:08, the instructor emphasizes "change," "sign," "row," and "column," which draws attention to the key aspects of the action he is performing.
  - At 13:53, the instructor emphasizes "define" and "action," indicating that he is identifying—and creating through this action—an important mathematical idea.
- *Mathematical speech*, which is the ways mathematicians use particular words or phrases (e.g. the meaning of "we" in mathematical discourse). For example:
  - At 12:17, the instructor says: "You recall—so I gave a proof last time." At 12:30, the instructor says: "You'll recall the general idea was this." Later, at 12:39, the instructor says: "Let's recall what a determinant is." In all of these cases, the implied observer recognizes "recall" as a cue to call to mind the relevant prior experience/idea, although the ways that these ideas will be used in the lecture are different. In the former case, recalling the proof provides continuity between points in time, while in the latter the definition—and related ideas and procedures—are used to construct an argument.
- *Gestures*, which can direct focus , depict the instantiation and action upon mathematical objects, or add additional meaning to some types of writing. For example:
  - At 12:27, the instructor holds his hand under the phrase "by determinants" while saying "And I'm just going to let this stand." The implied observer interprets this pointing gesture as the pointing to the argument that is represented by the phrase.
  - At 12:37, the instructor verbally refers to a matrix and, as he does so, holds up two hands with his index fingers pointing upward. The implied observer interprets this representation gesture as the instantiation of a matrix, with the fingers representing the vertical lines that enclose it.
  - At 13:53, the instructor describes the "action of a group on a set" and, while doing so, holds out one hand cupped up, then flips it over while he moves his arm in a semi-circle. The implied observer interprets this as performing an action performed by elements of a group.

Competencies.

While codes enable the implied observer to establish a mathematical and didactical context within which to operate, the *competencies* of the implied observer are the knowledge, skills, and understandings that are required for the observer to understand and work within the context. Some of these competencies may stem from the course in which the lecture is conducted, while other competencies may be drawn from other courses or knowledge that has been developed as a result of taking multiple mathematics courses (such as knowledge of mathematical processes like problem solving).

- *Knowledge of mathematical definitions and concepts from within the course*
  - Ring element (a, b, c, d)
  - Cycles ()
  - Groups
  - Group action
  - Symmetric group (made up of permutations)
- *Knowledge of mathematical definitions and concepts from outside the course*. For example,
  - 12:37 "have a matrix"; also rows and columns
  - 12:43 determinant
  - 13:08 "sign" of a row
  - negative sign: 12:43 subtraction in ad-bc; 13:08 additive inverse in –a and –b; 13:32 "negative of"
  - Sets (of matrices)
- *Knowledge and skills with mathematical algorithms and processes (including symbol manipulation) from within the course*
  - Multiplying matrices by cycles
- *Knowledge and skills with mathematical algorithms and processes (including symbol manipulation) from outside the course*
  - 12:43 computing the determinant
- *The ability to treat mathematical concepts as objects*
  - Determinants are values that can have additive inverses (e.g. it's not the process of computing ad-bc that's important)
  - Groups are sets (rather than processes of creating the sets) that can be acted upon
  - Matrices are objects that can be acted upon
  - Groups are entities that can act upon other sets
- *The ability to instantiate a specific example from a general definition or theorem (i.e. to "apply" definitions and theorems to create examples)*
  - [a, b, c, d] is a generic—yet specific—matrix
  - (2514) is an element of a symmetric group
- *The ability to generalize an abstract definition or notice abstract patterns from/in a specific example*
  - Interchanging rows changes the sign of a determinant
  - Group action from (2514)M
- *Knowledge and understanding of mathematical problem-solving*
- *The ability to connect various levels of rigor to requirements for justification*
  - Examples—one well-chosen example is sufficient for a non-rigorous proof

In addition, there are specific competencies that are required to understand mathematical proofs. Mejia-Ramos et al.'s (2010) framework for proof-comprehension lists five additional competencies:

- *The knowledge and understanding of the meaning of terms and statements (including the meaning of the theorem, of the individual statements in the proof, and the meaning of terms in the proof)*
- *The skill to justify individual claims*

- *The ability to understand the logical structure of the proof*
- *The knowledge and understanding of higher-level ideas that provide form and direction to the proof*
- *The knowledge and understanding of the general method used by the proof*

Although the excerpt above does not contain a formal proof, it does include a reference to a previous proof (indicated with "By determinants"). This is an example of the implied observer understanding that a proof can have a "theme"—a higher-level idea—that constituted the general direction of the proof. The excerpt also includes an informal argument in which a generic 2x2 matrix is used to justify the claim that interchanging rows will change the sign of the determinant. In this case, the implied observer

- Understands the terms and their logical relation in the implicit if [you interchange rows]… then [the determinant changes sign] argument. In addition, the implied observer understands each computational step in the proof.
- Recognizes that cb-ad is the additive inverse of ad-bc
- Recognizes that the 2x2 matrix is a generic—yet at the same time, specific—example
- Understands how a computation can be used to justify a claim
- Understands how a generic example can be used to justify the statement in general

Behaviors.

Behaviors are actions—usually mental actions—that the implied observer takes. These include:

- *Distinguishing between mathematical and non-mathematical aspects of the lecture*. There may be many components of a lecture that are not directly related to the pedagogy or mathematical content. For example:
    - The instructor may pause or emphasize words or include verbal "flourishes" without connecting these actions to emphasize or create mathematical meaning. These variations in the spoken component need to be separated from the emphasis that is intended to highlight or emphasize particular mathematical ideas
    - Similarly, many of the instructor's gestures may simply be the result of somebody who "talks with their hands" rather than instances where the instructor is gesturing in order to draw attention to, instantiate, or act upon mathematical objects.
- *Monitoring personal understanding*. Throughout the lecture, the implied observer determines whether the lecture components are sufficient to understand the concepts and takes appropriate actions—such as thinking of additional examples or related ideas from the class—that can be used to support the development of this understanding.
- *Identifying "ideas" and "concepts" so they can be acted upon as objects.* There were numerous mathematical concepts used in the example above—such as matrices, groups, sets, symmetric groups, and permutations/cycles—that could (and did) need to be treated as objects in order to successfully apply the relevant competencies.
- *Recalling examples, proofs, definitions, theorems, or proof structures*. As some examples:
    - At 12:17, the instructor says: "You recall—so I gave a proof last time."
    - At 12:30, the instructor says: "You'll recall the general idea was this."
    - At 12:39, the instructor says: "Let's recall what a determinant is."

    - At 13:51, the instructor says: "And so, you'll recall the idea was to think …" In order to understand the subsequent claims and arguments, the implied observer attempts to recall these mathematical ideas.
- *Seeking out abstract structures or patterns in exemplars and connecting these to definitions and theorems*. For example:
    - The instructor presented an example of interchanging two rows in a (generic) 2x2 matrix and the resulting determinant changing signs. The implied observer looks for the general pattern—the way changing the matrix entries causes the sign change—in order to understand the proof.
    - The instructor uses the term "action" but only provides an example of a cycle acting on a matrix rather than a full definition. The implied reader identifies the salient features of this example to construct an interpretation of what a group action is.
- *Creating examples based on abstract definitions or theorems.* For example:
    - In order to flesh out the meaning of a proof being "by determinants," the implied observer—along with the instructor—constructs an example of a matrix.
- *Recognizing and keeping track of the macro- and micro-structure of the lecture*. In addition to having the codes for formatting and organizational cues, the implied observer seeks out aspects of the lecture that indicate they should attempt to apply these codes. In addition, the implied observer identifies the various lecture components and identifies how they are related to each other in order to follow arguments and make connections between ideas.
- *Engaging with ideas in the appropriate order*. When subsequent ideas build upon each other, it is necessary to think about them in a specific order. However, the spatial-temporal nature of the lecture sometimes dictates that relevant ideas need to be recalled from the past or introduced as an aside (e.g. as indicated by the instructor using a different part of the board).
- *Committing definitions (and other facts) to memory*. Since mathematical ideas build upon and relate to each other, the instructor often asks observers to instantly recall facts, which necessitates that these facts are memorized when they are initially presented.
- *Seeking out the method, main ideas, and rigor of mathematical justification*. In order to apply the relevant competencies to understand mathematical proofs, the implied observer attempts to identify main features of proofs.
- *Applying a critical and skeptical lens to mathematical claims—including comparing these claims to what is already known and making conjectures along with the instructor*. Engaging in the process of mathematical justification—and understanding why and how arguments are used to justify a claim—requires skepticism on the part of any observer (or participant) in mathematics. For example, the instructor claims that interchanging rows of a matrix changes the sign of its determinant and subsequently begins an informal justification. Without skepticism of this initial claim, an observer would not be able to understand how the logic of the justification connects to the initial statement.

Coordinating Aspects of the implied observer.

Not only does the implied observer possess these codes, competencies, and behaviors, but the implied observer also coordinates them in order to make sense of the mathematical ideas in the lecture.

As a first example, consider the instructor's gestures. The implied observer has a behavior to distinguish between gestures that are mathematical and those that are simply "talking with your hands." After making this distinction, the implied observer uses relevant codes to establish the mathematical context, and then uses relevant competencies to work within that context. For example, at 12:37 the instructor holds up his hands with vertical fingers as he describes a matrix. The implied observer identifies this gesture as mathematical, understand it as instantiating a matrix, and then thinks of the matrix as an object that has properties and can be acted upon.

As a second example, consider the instances where definitions and theorems and previous results are used without explicit instructions to "recall" them. The implied observer has a behavior to seek out these definitions and theorems and to recall these examples, codes to identify when they are actually being invoked, and relevant competencies to understand the way in which they are being used. For example, at the end of the excerpt, the instructor identifies "(2514)M" as an example of a cycle acting on a matrix. The implied observer identifies (2514), M as exemplars of definitions and the example as one that has previously been discussed. Then, the implied observer interprets (2514) as a cycle and M as a (5x5) matrix, and uses the relevant competencies to understand how to interpret this configuration of symbols and how to use this exemplar to understand the underlying abstract concept of a group action.

Discussion.

This paper makes one major and one minor contribution to the study of undergraduate teaching. The major contribution that this paper makes is to adapt Weinberg and Wiesner's (2011) framework for studying written mathematical text to create a framework for studying lectures. We have described the construct of an *implied observer* of a lecture that similarly consists of a bundle of codes, competencies and behaviors that are needed to respond to the lecture in a way that is meaningful and accurate.

In the process of adapting Weinberg and Wiesner's (2011) framework, we devised ways of capturing the "facts" of a lecture and a means of interpreting those facts. While there are models for capturing the elements of a lecture (such as Schoenfeld, 1999), they were adapted for goals other than describing the types of codes, competencies and behaviors of the implied observer. In order to describe the implied observer, we first described the components of the lecture as those things that can be directly observed; the written text, the gestures, and the spoken text. Yet, unlike a mathematics text where all the information is printed upon an unchanging page, a lecture includes both semi-permanent components such as writing and fleeting elements of the lecture that include gestures and speech. As a result, an additional component of the lecture is the temporal-spatial element. The temporal-spatial element includes describing when the instructor writes a particular piece of text on the board, where the instructor writes it, what components of the lecture came just before, what other components are concurrent, and what components comes just after. Much like the way the vertical and horizontal arrangement of words on a page mediates meaning in a text, the temporal-spatial component of a lecture has significant impact on the implied observer and the types of meanings that actual observers might take from the class.

In order to describe the codes, competencies and behaviors that comprise the implied observer, we included in our framework a method for describing the communicational and mathematical *aspects* of the class. The communicational aspect principally focuses on organizational cues that alert the observer to the structure of the lesson. These might include written outlines, lines drawn on the board as organizers of board space, or spoken cues that describe the goals of a particular part of class. The mathematical aspect of the framework includes facts, procedures and algorithms, as well as process such as problem-solving, communication, representation and justification. While not important in the section of class presented here, the nature of an abstract algebra class suggests that extra attention to the codes, competencies and behaviors needed to comprehend proof is important and will be included in future work. We demonstrated that this framework could be applied to analyze a proof-based class and give a meaningful interpretation of the lecture.

The minor contribution of this paper is represented by the analysis of three minutes of an abstract algebra class that we carried out. "Very little empirical research has yet described and analyzed the practices of teachers of mathematics" (Speer, et al., 2010, p. 99) at the undergraduate level despite repeated suggestions for this type of study (Harel & Sowder, 2007; Harel & Fuller, 2009, Speer, et al. 2010). We presented a single, exploratory case study focused on a few minutes of instruction, and recognize it is inappropriate to draw generalizations from it. Yet, even this represents a small contribution to the research literature due to the need for a body of empirical evidence to provide a basis for more theoretical work. The second aspect of this contribution is the creation of the description of the implied observer. Our analysis showed that the implied observer of this section of the abstract algebra class under study included a complex constellation of codes, competencies, and behaviors.

Some important codes that the implied observer possess are those related to a combination of organizational cues and symbols. For the implied observer to be able to correctly interpret the instructor's use of formatting of the board, using the far right column for reminders about ideas related to the proof, but not directly required by the proof, also requires the implied observer to have codes that mark transitions between class topics. It is only because no such codes were invoked when the lecturer moved to work on the right side of the board that the implied observer would interpret the coming material to be related to the proof. The codes for symbols are invoked any time that the lecturer is writing on the board, and, as a result, for the implied observer to make sense of that text, some of which is distinct from his speech.

Perhaps the single most important competency that the implied observer possess is the ability to act upon mathematical concepts as objects. Throughout this portion of class the lecture requires operation on determinants, matrices, and groups and finally, groups are used to act upon matrices. A second important pair of competencies is the ability to instantiate a definition, procedure, or theorem via a specific example, or to generalize from a specific to a definition, procedure or theorem. The lecture requires this in multiple instanced related to matrices, cycles, permutations, and group actions. In fact, nearly all of the mathematical content of this section of class draws upon this competency.

Finally, one of the most important behaviors of the implied observer is to distinguish between mathematical and non-mathematical aspects of the lecture. The lecturer "talks with his hands," pauses, emphasizes words, repeats phrases, and includes asides (although not observed in this section of class). In short, when his corpus of words, writing and gestures are taken together they are continuous, and include nearly overwhelming possibilities for communication. The ability to filter out the 'signal' from the 'noise' is of critical importance to making meaning from the class.

Fundamentally, this framework and mode of analysis is still quite preliminary and there are significant theoretical and methodological questions that remain, even before it can be used as a means to describe the opportunity to learn and contrast the profile of the implied observer with actual observers. The remainder of this section will first describe open theoretical questions and their implications for the framework and methods. Subsequently, it will describe issues and directions for the future that relate to the codes, competencies, behaviors and learning of actual students.

One of the most important theoretical questions that remains is whether the lecture should be understood as a *closed* text—one that has a single meaningful interpretation—or an *open* text that has multiple meaningful interpretations. While the instructor of the course does have a particular understanding of the content that he is likely attempting to impart, we have not yet decided whether an observer of the lecture might develop meaningful mathematical understandings that differ significantly from the lecturer's intentions but would still be accepted as correct or valid by the mathematics community. There are then multiple issues that stem from this question. For example, interpreting the lecture as a closed text might imply that the implied observer is comprised of a *unique* set of codes, competencies and behaviors and that there is a maximal amount of meaning that any observer might take from a particular lecture. Yet, it is not clear that there is a maximal amount of meaning that a collection of even highly trained mathematicians could agree upon due to the fact that each might see echoes of more advanced topics that the implied observer could be given access to via the lecture. Would we then attempt to describe the union of all possible meanings? The intersection? Or some other combination?

If we treat the lecture as an open text that has multiple mathematically meaningful interpretations, then describing the implied observer becomes more problematic in that there might be several possible collections of codes, competencies and behaviors that give access to the same meaning, or there may be different meanings that can be understood as appropriate and reasonable. If there are multiple such possible meanings, it then requires standards to judge the appropriateness or reasonableness of the meaning that might be made. Similarly, we might ask whether we would attempt to describe the union or intersection of the meanings, and, if we did, what insight into the implied observers that would give us.

Even the excerpt of the class that we presented methodological difficulties related to interpretation of the codes the implied observer possesses and the meaning that the implied observer makes. The three principle authors disagreed about whether a gesture that the instructor made carried mathematical meaning. At 12:39, the instructor said: "let's recall what a determinant is" while moving his left hand in a small circle in the space that had previously been indicated as a matrix by both hands. The disagreement concerned whether the gesture communicated that the determinant was computed via operations on the entries of the matrix, or, whether the gesture indicated that the observers should engage in recalling the fact—a gesture that doesn't have clear mathematical meaning. Yet all of us agreed on the mathematical knowledge that was being prompted for recall regardless of the interpretation of the gesture.

One of the questions that bridges the theoretical and methodological is the issue of the temporal nature of the lecture and the fact that spoken and gestural elements are only accessible in-the-moment—and only if they are noticed by the (actual) observer. Thus, we have hypothesized that the implied observer may have a behavior of never being distracted, and is able to simultaneously focus on the instructor's presentation while also having nearly instantaneous checking for mathematical meaning and meta-cognitive processing. We relate this to van Es and Sherin's (2008) work on noticing, and the fact that novices in a particular situation cannot notice as much about the situation as an expert. As a result, it is arguable that the implied observer would need to be infinitely capable of noticing components of class and apprehending them. The implication is that otherwise the implied observer would possibly miss something that happens due to the impermanence of speech and gesture, yet that differs significantly from the implied reader and we have not managed to resolve this difficulty. In a similar way, we have not decided whether the implied observer would take notes. The instructor behaves in a way that is designed to allow students to take notes (and he has even told students that they should be taking notes) but we cannot decide whether taking notes is an essential part of the meaning-making process.

Once we resolve the theoretical and methodological issues, we intend to expand our study to include students (i.e. the actual observers) of the lesson and what they learn from a class (at least short-term learning). In order to understand how undergraduate mathematics students' learn class material, it is important for us to describe and better understand their in-class experiences and how they construct meaning from mathematics lectures. In particular, we view students' opportunity to learn mathematics from a lecture as the interface between the implied observer and the actual observer—a student's own behaviors, codes, and competencies, along with the ways they comprehend proofs. Moreover, in describing the actual observer we will have to further wrestle with the temporal nature of speech and gesture as well as the student's ability to notice. Although the framework described here has significant questions remaining, we hope that it provides a valuable theoretical lens for beginning to describe the difficulties students may have making sense of lectures and, by doing so, help students engage with and learn from mathematics lectures.

References

Armbruster, B. B. (2000). Taking notes from lectures. In R. F. Flippo & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research.* Hillsdale, NJ: Eribaum.

Fukawa-Connelly, T. (in press). A case study of one instructor's lecture-based teaching of proof in abstract algebra. *Educational Studies in Mathematics*.

Harel, G., & Fuller, E. (2009). Contributions toward perspectives on learning and teaching proof. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 355-370). New York, NY: Routledge.

Harel, G., & Sowder, L. (2007). Towards a comprehensive perspective on proof. In F. Lester (Ed.), *Second handbook of research on mathematical teaching and learning* (pp. 805-842). Washington, DC: NCTM.

Mehrabian, A. (1971) *Silent Messages*. Belmont, CA: Wadsworth.

Mejia-Ramos, J., Weber, K., Fuller, E., Samkoff, A., Search, R., Rhoads, K. (2010). Modeling the comprehension of proofs in undergraduate mathematics. *Proceedings of the Conference on Research in Undergraduate Mathematics Education,* Raleigh, NC. http://sigmaa.maa.org/rume/crume2010/Archive/Mejia-Ramos%20et%20al..pdf

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.

Shein, P. (2012). Seeing with two eyes: A teacher's use of gestures in questioning and revoicing to engage english language learners in the repair of mathematical errors. *Journal for Research in Mathematics Education*, *43*, 182-222.

Schoenfeld, A. H. (1999). Models of the teaching process. *Journal of Mathematical Behavior, 18*, 243-261.

Speer, N., Smith, J., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99-114.

van Es, E. A. & Sherin, M. G. (2008). Mathematics teachers "learning to notice" in the context of a video club. *Teaching and Teacher Education*, 24, 244-276.

Weinberg, A. & Wiesner, E. (2011) Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics 76* (1), 49-63.

UNDERGRADUATE PROOF IN THE CONTEXT OF INQUIRY-BASED LEARNING

Dr. Todd A. Grundmeier, Alyssa Eubank, Shawn Garrity,
Alyssa N. Hamlin, and Dr. Dylan Retsek
California Polytechnic State University, San Luis Obispo

Students' proof abilities were explored in the context of an inquiry-based learning (IBL) approach to teaching an introductory proofs course. IBL is a teaching method that focuses on student discussion and exploration in contrast to lecture-based instruction. Data was collected from three sections of an introductory proofs course, which included 70 students total. Data collection included a portfolio from each student, consisting of their work on every proof assigned throughout the course, as well as each student's final exam. Contrary to previously published research related to courses taught in a more traditional lecture-based setting, this data analysis suggests that students developed an understanding of how to correctly use definitions and assumptions within the context of their proofs. Results also suggest that within the IBL setting, students generally organized their proofs in an efficient, thoughtful, and logical manner.

Key-Words: Proof, Inquiry-Based Learning, Definitions, Assumptions, Structure

Introduction

Mathematics instructors always face the challenge of teaching in a manner that allows students to develop conceptual understanding of material. Current undergraduate teaching typically consist of in-class lecture followed by students completing homework outside of class. This classroom structure may not encourage the development of deep problem solving techniques that students will be able to utilize in higher-level classes. After observing this weakness in our current system, educators have reevaluated their methods. An emerging method to combat these potential problems is inquiry-based learning (IBL).

Stemming from the Modified Moore Method (MMM), IBL is a teaching method that focuses on student discussion and exploration in contrast to lecture-based instruction. According to Schinck (2011), the MMM:

> … [has] students pose conjectures, construct their own proofs, justify their reasoning to their peers at the board, and assess the validity of proposed solutions and proofs. Textbooks [are] generally not used. Lectures [are] kept to a minimum… Student collaboration is sometimes encouraged, with solutions to problems shared during small group and/or whole-group discussions.

The degree to which these guidelines are implemented within an IBL classroom varies from teacher to teacher; more often than not, you will find the instructor interacting with the class in some way. Typically, instructors place a high responsibility on students for their own learning and use leading questions to guide them.

"As mathematics education researchers turn their attention to IBL, evidence mounts that this approach to the teaching of mathematics is ideal for the teaching of proof" (Schinck 2011). Studies summarized in Schinck's (2011) article conducted by Boaler (1998) and Rasmussen and Kwon (2007) deduce that IBL students experience mathematics in a way that deepens their comprehension of abstract ideas essential to proof. IBL students also seem to be more creative and better prepared to utilize knowledge to solve new problem types compared to students in traditional teaching methods.

The data collected for this exploration are from three sections of a Methods of Proof course taught using an IBL style. This sophomore-level course is designed as an introduction to formal notation and techniques essential for the development of logical mathematical proof. This paper will discuss the student work that was a product of the course and will explore the possible effects the IBL teaching method had on student understanding and proof technique. First it is important to set the context in relation to previous research on mathematical proof.

Classification of Proof Schemes

High school students typically acquire little understanding of mathematical proof; hence, universities must emphasize instruction geared towards gaining maximum student understanding in introductory proof classes (Harel & Sowder, 1998). In order to best support students, instructors must first be able to identify the ways students think and the proof schemes they use. Harel and Sowder (1998) define proof scheme as "what convinces a person, and … what the person offers to convince others." Knowing that students' proof schemes allow instructors to identify student understanding, Harel and Sowder (1998) devised three categories to describe these schemes: external conviction, empirical, and analytical. Each describes a different level of mathematical formality and shows various levels of critical thinking.

The external conviction proof scheme classifies student proofs that depend on outside sources and reflect little independent thought (Harel & Sowder, 1998). For the empirical proof scheme, all cases of the problem are not addressed or not sufficiently justified; hence this proof scheme does not represent convincing or rigorous mathematical proof. Finally, the analytical proof scheme "is one that validates conjectures by means of logical deductions" (Harel & Sowder, 1998). The latter is considered to be the most rigorous proof scheme and consequently instructors must hold students to this caliber of proof production.

Others have also categorized the techniques that students use to convince themselves and others of mathematical statements. Weber (2005) explores "the relationship between problem-solving processes and learning opportunities in the activity of proof construction". He placed the various types of proof productions that students use into three categories: procedural, syntactic, and semantic. A procedural proof relies on a similar proof as a template. A template is often a proof that the instructor demonstrates in class, but can also come from a textbook or other source of authority. A syntactic proof relies on the student's ability to use theorems and logical rules in an effective way. Writing syntactic proofs amounts to "logically manipulating mathematical statements without referring to intuitive representations of mathematical concepts" (Weber 2005). Both syntactic and semantic proof styles play a role in the development of student mathematical thinking.

In their 2009 paper, Weber and Mejia-Ramos respond to a previous study by Alcock and Inglis (2009) in which they describe a new way to define syntactic and semantic, and thus categorize proofs differently. However, for Weber and Mejia-Ramos, it is not enough to label student proofs as using syntactic or semantic reasoning without investigating the ways in which the different reasoning types affect proof construction. Thus Weber and Mejia-Ramos prefer to consider *how* semantic and syntactic reasoning are used as opposed to simply *if* they are used (Weber & Mejia-Ramos 2009).

Analyzing Student Proof

To learn more about student understandings of formal mathematical proof, researchers have created coding schemes to analyze student work. Moore (1994) studied and analyzed student difficulties in an introductory proof class. In his study, the course observed was taught with the traditional lecture-based method. "The data were collected primarily through nonparticipant observation of class each day, interviews with the professor and the students, and tutorial sessions with the students outside of class" (Moore 1994). Upon analysis, Moore found seven major problem areas:

D1. The students did not know the definitions, that is, they were unable to state the definitions.
D2. The students had little intuitive understanding of the concepts.
D3. The students' concept images were inadequate for doing the proofs.
D4. The students were unable, or unwilling, to generate and use their own examples.
D5. The students did not know how to use definitions to obtain the overall structure of proofs.
D6. The students were unable to understand and use mathematical language and notation.
D7. The students did not know how to begin proofs.

The way in which Moore (1994) collected data enabled him to think from the perspective of the professor and students. He acknowledges the likelihood that "a difficulty or lack of understanding in one area led to further difficulties in another area." Moore (1994) concludes that these seven problem areas "were cognitive and [students] would have encountered these difficulties despite diligent studying."

Weber (2010) continues his investigations of student thought processes during proof construction, this time focusing on what makes a proof convincing to students. He discovered that arguments students find convincing are not necessarily those that they consider valid proofs, and vice versa. In fact, many students (and surprisingly some math teachers as well) "believe the format of an argument is more important than its content in judging whether an argument is a proof" (Weber 2010). This leads to students considering certain arguments to be correct proofs despite obvious logical flaws simply because they are structured like the proofs they have seen in class. Such inability to judge the validity of arguments will make it difficult for students to understand proofs in upper division math classes. Moreover, students are often convinced by empirical arguments even though they realize that these arguments do not prove the statement. As referenced by Weber (2010), "Healy and Hoyles (2000) found that many high school students who personally prefer empirical arguments recognize that they would not receive high marks from their teacher." Thus while students do have some understanding about what is expected in the mathematical community, this understanding does not necessarily translate to their proof techniques.

There are also inconsistencies between what instructors expect from students and what the students think they are being asked to do. These inconsistencies lead to student difficulties when constructing proofs. In order to better align student and instructor expectations, Andrew (2009) created his own proof error evaluation tool (PEET). According to Andrew (2009), for a PEET to effectively address inconsistencies, it must take into consideration instructors' varying beliefs and expectations about which parts of proof construction are most important. At the same time it must also be understandable to students, as students often adjust their work to match the grading schemes they are given (Andrew 2009).

When creating the PEET, Andrew (2009) used a four-phase process beginning with his own perceptions of proof construction and difficulties from when he was a student. He then entered the second phase of the PEET creation, which incorporated ideas from textbooks and other sources. At this point in the process he had a plethora of ideas, which he grouped into categories in the third stage. In the fourth and final stage of the process, Andrew tested his compiled list against student-generated proofs and adjusted it to include other cases of mistakes or misconceptions that were not previously addressed. It was then tested on middle school and graduate level proofs, both student and researcher generated. The refined PEET has 22 codes, which address both proof structure and conceptual understanding.

Andrew has found that when used in a classroom, his developed coding system "supports students in becoming better proof writers because it encourages instructors to give students quality feedback of their proof-writing skills … [and] encourages students to reflect on their own proofs and learn from their mistakes" (Andrew 2009).

Procedure

Course, Coded Problems, and Coding Scheme

An adaptation of the IBL method was used to teach three sections of Methods of Proof, an introductory mathematical proofs course. During the course, the 70 students were required to submit homework problems as well as complete a portfolio consisting of all problems assigned. The two main class activities during the quarter were student presentation of proofs and group problem solving. Student presentation of proof accounted for approximately 75% of class time, and the majority of assigned problems were presented. Group problem solving only occurred when the entire class was at an impasse in the construction of a proof; this was the other 25% of class time. During presentations, the professor said as little as possible, occasionally asking leading questions to encourage student discussion in order for the class to come to a consensus on a correct proof.

For the purposes of this research, we chose ten problems to evaluate using a coding scheme developed from past work on mathematical proof. Due to some variation in presented problems among the three sections of the class, we set the goal of finding five problems to code that were presented in all three sections and five that were not. Four of the chosen problems presented in class were also turned in as homework, while the other six chosen problems were only included in the portfolio. When choosing the problems, we chose two comparable problems from each major content area of the course– one presented and one not. The table below includes the chosen problems, whether they were presented in class, and how they were collected.

Table 1: Summary of Coded Problems

2.17	Presented, Homework	Let a and b be positive integers. The integer $a+1$ divides b and the integer b divides $b+3$ if and only if $a=2$ and $b=3$.
2.18	Not Presented, Portfolio	Let x be a real number. The quadratic $x^2+2x+1=0$ if and only if $x=-1$.
3.24	Presented, Homework	Let A, B, C and D be sets. If $A\cup B\subseteq C\cup D$, $A\cap B=\phi$, and $C\subseteq A$, then $B\subseteq D$.
3.25	Not Presented, Portfolio	If A, B, and C are sets, then $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

3.35	Presented, Homework	Let $\mathbf{A} = \{(-a,a) \subseteq R : a > 0\}$. Determine $\bigcup_{A\in\mathbf{A}} A$ and $\bigcap_{A\in\mathbf{A}} A$.
3.49	Not Presented, Portfolio	If $\mathbf{A} = \{A_\alpha : \alpha \in \Delta\}$ is an indexed family of sets and B is a set, then $B \cup \left(\bigcup_{\alpha\in\Delta} A_\alpha\right) = \bigcap_{\alpha\in\Delta} (B \cup A_\alpha)$
4.23b	Presented, Homework	Let *A, B, C*, and *D* be sets. Let *R* be a relation from *A* to *B*, *S* a relation from *B* to *C*, and *T* a relation from *C* to *D*. Then $T \circ (S \circ R) = (T \circ S) \circ R$.
4.41a	Not Presented, Portfolio	Let $\mathbf{A}$ be a partition of the nonempty set *A*. Define the relation *Q* on *A* by $Q = \{(x,y) \in A \times A : (\exists C \in \mathbf{A})(x \in C \wedge y \in C)\}$. Then Q is an equivalence relation on *A*.
5.16	Presented, Portfolio	If $f : A \to B$ and $g : B \to C$, then $g \circ f : A \to C$.
5.19	Not Presented, Portfolio	Show that the inverse relation f^{-1} to the function $f : R \to R$ given by $f(x) = 2x^2 + 1$ is not itself a function.

To assess students' understanding of mathematical notation and proof techniques, we developed a coding scheme to evaluate the selected student work. The coding scheme used is an adaptation of work by Harel and Sowder (1998) as well as Andrew (2009).

1st Level Coding

All solutions to the ten problems collected were first coded based on attempted proof scheme. Each attempted proof was coded as having used an *analytical* proof scheme, an *empirical* proof scheme, or *other* proof scheme (not analytical or empirical). Harel and Sowder's (1998) definitions of analytical and empirical proof schemes were used for this level of coding. As the purpose of this research project is to analyze student thought processes when developing formal proofs, it was deemed necessary to further code proofs that were deemed analytical.

2nd Level Coding

The following codes for analytical proofs were adapted from Andrew (2009) to explore students' use of language and notation and to identify errors with implications or steps within a proof. In this coding, the S (structure) codes identify language and notation errors while U (understanding) codes identify errors in implications or steps.

S3 – Proof ideas not in logical order
S4 – Contains extra details or steps; unnecessarily long and hard to follow
S5 – Illegible or difficult to read
S8 – Nonstandard or confusing notation
U4 – A crucial step is not sufficiently justified; important parts of proof not addressed
U5 – False statement or incorrect computation; incorrect implication or equivalence
U6 – Forgot to include a non-trivial case; did not address one aspect of the problem
U7 – Forgot to include a conclusion for something proved

In addition to language, notation, and implication issues, we were interested in determining how well students apply definitions and make assumptions in the context of their proofs. Therefore the 2nd level of coding also involved recording the number of assumptions and definitions used in each analytical proof; it was then noted whether each assumption and definition was correct or incorrect.

Results

Analytical, Empirical, and Other Proofs

Students' proofs are conceived using different levels of deduction that result from their diverse thought processes. Research implies that the level of deduction a student uses in a proof can be classified (Harel and Sowder, 1998). The 1st level of coding, as previously described, was used to classify students' proofs in this manner. Of the 548 attempted proofs coded, 473 (86.31%) were coded analytical, 4.2% were coded empirical, and 9.49% were coded other.

For seven of the ten problems coded, more than 95% of student proofs were analytical, which implies that students composed formal proofs and attempted logical deduction. It is interesting to note that a smaller percentage of analytical proofs were recorded for problems 3.35 (27.59%), 4.41a (78.26%) and 5.19 (63.79%). Since the majority of the non-analytical proofs on problem 5.19 were empirical (32.76%), it was common for students to rely heavily on specific examples when trying to prove this problem. On the other hand, the majority of non-analytical proofs for problems 3.35 and 4.41a were recorded as *other* (67.24% and 21.74%) which means students did not attempt formal deduction or rely on specific examples. All results that follow only include the 473 proofs coded as *analytical*.

Assumptions and Definitions

Research suggests that students struggle with the use of assumptions and definitions when first introduced to mathematical proof (Moore, 1994). Students' use of definitions and assumptions was investigated with a particular focus on how often they incorrectly made assumptions and incorrectly used definitions and Table 2 provides an overview of this analysis.

Table 2: Definitions and Assumptions in All Problems

	Problem										
	2.17	***2.18***	***3.24***	***3.25***	***3.35***	***3.49***	***4.23b***	***4.41a***	***5.16***	***5.19***	***All***
# Coded	61	55	62	52	15	52	64	18	56	37	473
% Incorrect Definitions	0	0	2.55	1.77	0	4.14	2.88	8.75	1.78	3.78	2.57
% Incorrect Assumptions	0	1.93	0	7.32	0	3.19	1.05	18.18	9.09	0	3.53

The overall low incorrect percentages suggest that students consistently used definitions and assumptions appropriately. There is only slight variation in the percentages of incorrect definitions amongst the problems. However there is more variation in the incorrect use of assumptions. It is logical to explore if these differences are related to problem type or content area.

A key feature of the IBL Methods of Proof course was the responsibility of students to present proofs daily. Table 3 describes students' use of definitions and assumptions based on whether problems were presented in class.

Table 3: Definitions and Assumptions in Presented and Not Presented Problems

	# Coded	% Incorrect Definitions	% Incorrect Assumptions
Presented	259	2.03	1.71
Not Presented	214	3.36	6.5

When problems were not presented in class, students were slightly more likely to use definitions incorrectly. Students also committed approximately 5% more errors when making assumptions in proofs that were not presented.

The instructor made the decision to structure the class around problems collected for homework and problems completed for student portfolios. One might question if the requirement to hand a proof in for homework influenced how students worked on problems. Table 4 details the definition and assumption use based on whether problems were collected only as part of the portfolio or as homework.

Table 4: Definitions and Assumptions in Portfolio and Homework Problems

	# Coded	% Incorrect Definitions	% Incorrect Assumptions
Portfolio	270	2.89	7.07
Homework	203	2.13	0.16

Students used a slightly higher percentage of definitions incorrectly in portfolio problems. In addition, the percent of assumptions made incorrectly in the portfolio problems was about 7% more than that of homework problems.

The Methods of Proof class covered various fundamental ideas for upper division math classes. Within each content area, there reside unique conceptual challenges for introductory level students. Table 5 depicts the usage of definition and assumption by course content.

Table 5: Definitions and Assumptions in Course Content

	Content Area				
	Set Equality	***Indexed Families***	***Union and Intersection***	***Relations***	***Functions***
# Coded	240	68	182	82	93
% Incorrect Definitions	2.56	4.07	2.71	3.77	2.01
% Incorrect Assumptions	5.30	3.15	2.68	8.72	9.02

The highest percent of incorrect definitions occurs within the Indexed Families content area. Relations and Functions hold the highest percent of incorrect assumptions. The lowest percent of incorrect assumptions come from problems related to Union and Intersection. In general the results are similar except when examining the use of assumptions for problems related to Relations and Functions. These problems are the last two chapters of the course work and about 9% of student assumptions were incorrect on these problems.

S and U Codes

The use of mathematical language and notation as well as the ability to correctly use and justify implications and steps within a proof are important skills to develop in an introductory proof class. Recall that S codes identify errors related to the structure of the proof and U codes identify errors relating to student understanding.

After coding Problem 3.35 it seemed necessary to look at the data closely. While its distribution of S codes seems to fall within the range of the other problems, it stands out as the only problem with 100% of the coded proofs having more than two U codes; in fact all students had at least six U codes. The average number of U codes per problem is more than 10, whereas the next highest average is only 2.6 codes per problem. The vast number of U codes affected most of the groupings of coded problems, skewing the data to suggest that whichever group it was a part of had a much higher number of U codes. Problem 3.35 was a two-part problem where both parts required biconditional proofs. It was common for students to completely forget a part of the problem, which resulted in a U6 for each direction and case within the part they left out. While this large number of U codes was deserved, it perhaps overstated the problems students were having with the content. Due to these issues data analysis for S and U codes was calculated both including and excluding problem 3.35. This analysis made it clear 3.35 was enough of an outlier to be excluded from the data and discussion of S and U codes. Table 6 provides S and U data for presented proofs, not presented proofs and all proofs coded analytical.

Table 6: Average S and U codes per problem and overall

	Coded	**S3**	**S4**	**S5**	**S8**	**Avg S**	**U4**	**U5**	**U6**	**U7**	**Avg U**
Presented	243	.21	.22	0	.39	.82	.60	.41	.32	.26	1.59
Not Presented	214	.04	.15	0	.44	.63	.63	.24	.10	.15	1.12
All	**457**	**.13**	**.18**	**.005**	**.41**	**.72**	**.62**	**.33**	**.22**	**.21**	**1.38**

Focusing on the overall averages from Table 6, the most common code given was U4 signifying incomplete justifications for proof steps. Approximately six out of every ten problems earned this code and it makes up for nearly half of the total U code average. Additionally, it is clear that illegibility of proofs is not an issue since only one of the 457 coded problems received an S5 code. S3 and S4 codes, denoting illogical proof order and extra details or steps respectively, appear in approximately 2 out of every 10 problems. Overall, there are almost twice as many total U codes as total S codes, meaning that the bulk of student error does not lie with proof structure but with understanding proof techniques.

The course focus on student presentation again raises the question of whether the characteristics of proof were similar for problems that were and were not presented in class. It seems that students had fewer issues with language and notation (denoted by the S codes) and understanding (denoted by U codes) on problems that were not presented in class. In both presented and not presented problems, the most common S code was S8, given when students use confusing notation or wording. This was the only instance with more S codes on not presented problems. However, the low occurrence of S3 codes given when proof ideas are not presented in a logical order, suggests that students are relatively competent in organizing their thoughts even though they may struggle with expressing them in understandable ways.

The most common U code in both categories was U4, which is given when students do not completely justify a step; this was also the only instance with more U codes on not presented problems. For both presented and not presented problems, students received on average almost two times as many U codes as S codes. This again suggests that there are more problems with understanding of implications within proofs than there are with language and notation errors.

Another aspect of the course was that only a selected number of problems were turned in and graded for homework. Table 7 addresses this relationship by showing the S and U code data for homework problems and those found only in the student portfolio.

Table 7: Average S and U codes per problem by portfolio and homework problems

	Coded	**S3**	**S4**	**S5**	**S8**	**Avg S**	**U4**	**U5**	**U6**	**U7**	**Avg U**
Homework	187	.26	.27	.01	.31	.85	.46	.41	.21	.19	1.27
Portfolio	270	.04	.13	0	.49	.66	.73	.27	.22	.22	1.44

There does not appear to be much contrast for U and S codes when comparing homework to portfolio. S8 is the only S code that is more common in the portfolio than in the homework. Approximately one in every two problems from the portfolio received an S8 code, denoting confusing notation or wording, and on average a little more than three out of every 10 homework problems received the same code. Some errors were more prevalent on homework. For example, almost one in every four homework problems received an S3 code (given when proof ideas are not presented in a logical order) whereas less than one in 25 of the portfolio problems received this error. U4 codes were common in both categories, appearing in approximately two out of every four proofs from the Homework and three out of four proofs in the Portfolio.

Similar to the data organization for definitions and assumptions, student errors in structure and understanding were explored in relation to the different content areas of the course. Table 8 shows this distribution.

Table 8: Average S and U codes per problem by content area

	Coded	**S3**	**S4**	**S5**	**S8**	**Avg S**	**U4**	**U5**	**U6**	**U7**	**Avg U**
Set Equality	224	.23	.27	.01	.67	1.18	.87	.40	.30	.30	1.87
Indexed Families	52	.10	.31	0	.52	.93	1.27	.27	.15	.27	1.96
Union & Intersection	166	.07	.19	0	.42	.68	.77	.26	.11	.14	1.28
Relations	82	.54	.49	.01	.67	1.71	.57	.57	.16	.31	1.61
Functions	93	.01	.07	0	.52	.60	.76	.42	.42	.29	1.89

Proofs related to Relations hold the highest average for every S code. It is clear that students have more language and notation errors in the Relations content area than any other. Specifically for average S3 codes, Relations have more than double the amount of errors than the next highest content area. Proofs in Relations also have the highest average of U5 codes, which means more incorrect implications occurred in this particular content area than the others. S8 and U4 are the most popular S and U codes in every content area. This implies students struggled with formal notation and sufficiently justifying all steps of their proofs. The Unions and Intersections content area seems to stand out for having a U7 average that is half of the average of the rest of the areas. Thus, proofs in the Unions and Intersections category had formal conclusions more than twice as often as any other category.

Proofs related to Functions and Unions and Intersections seemed to give students the least amount of trouble with language and notation errors as evidenced by the lower number of S codes. Proofs related to Unions and Intersections contained the fewest errors in implications or steps, as evidenced by the low number of U codes.

Final

Two slightly different versions of the final exam were given to the three sections of the class. The exam structure included three distinct parts. The first included problems that were part of previous coursework that students had already completed for their portfolio. The second required proving new results using definitions and theorems that were familiar from class. The third required proving theorems related to previously unseen definitions. The exams provided the opportunity to examine how the IBL experience translated to using new definitions, making assumptions and proving previously unseen results in a time-constrained situation. Therefore, the previously described coding scheme was applied to the two problems from the third section of the final exams. The definitions and propositions and the results related to the coding of the four new propositions follow.

Definition: Let $K \subseteq R$. The point $p \in R$ is called a *limit point* for K if for every $\delta > 0$ there exist at least two elements of K in the open interval $(p-\delta, p+\delta)$.
Definition: Let $K \subseteq R$. The set K is called *closed* if every limit point of K is an element of K.
Proposition 1.1: If $\{K_\alpha\}_{\alpha\in\Delta}$ is an indexed family of closed sets, then $\bigcap_{\alpha\in\Delta} K_\alpha$ is closed.
Definition: Let $E \subseteq R$. The set E is called *open* if its complement $\overline{E}$ is closed.
Proposition 1.2: If $\{E_\alpha\}_{\alpha\in\Delta}$ is an indexed family of open sets, then $\bigcup_{\alpha\in\Delta} E_\alpha$ is open.

Definition: Let $E \subseteq R$. The set E is called *open* if for every $p \in E$ there exists $\delta > 0$ such that the open interval $(p-\delta, p+\delta)$ is contained in E.
Proposition 2.1: If $\{E_\alpha\}_{\alpha\in\Delta}$ is an indexed family of open sets, then $\bigcup_{\alpha\in\Delta} E_\alpha$ is open.
Definition: Let $K \subseteq R$. The set K is called *closed* if its complement $\overline{K}$ is open.
Proposition 2.2: If $\{K_\alpha\}_{\alpha\in\Delta}$ is an indexed family of closed sets, then $\bigcap_{\alpha\in\Delta} K_\alpha$ is closed.

The initial coding of the exam problems was to classify student proof attempts as analytical, empirical, or other. Since this was an exam, blank problems existed and it was decided to code these as other. Similarly to our findings for the coursework proofs, the majority of proofs coded on the final exam were analytical (70%). Empirical and other codes each accounted for 15% of the attempted solutions. The problems deemed *analytical* were evaluated with the 2nd level of coding. The use of definitions and assumptions on these four problems is highlighted in Table 9.

Table 9: Definitions and assumptions on final exam problems

	Problem				
	1.1	***1.2***	***2.1***	***2.2***	***All***
# Coded	34	28	18	15	95
% Incorrect Definitions	11.49	15.94	15.39	10.35	13.3
% Incorrect Assumptions	3.45	0	8.00	0	2.96

On average, students used definitions incorrectly 13.3% of the time while trying to write proofs on the final exams. On two of the four problems, students did not make any incorrect assumptions and overall only made incorrect assumptions 2.96% of the time.

Table 10 shows the distribution of S and U codes on the problems from the final exams that were coded as analytical.

Table 10: Average S and U codes per problem on final exams

	Coded	**S3**	**S4**	**S5**	**S8**	**Avg S**	**U4**	**U5**	**U6**	**U7**	**Avg U**
All	95	.08	.31	0	.56	.95	.70	.55	.31	.06	1.62

Since U codes were more prominent than S codes on the final exams, it seems students had more issues with errors related to understanding within their attempted proofs than they did with issues of language and notation. This becomes clear in the fact that there were 1.56 total U4, U5 and U6 codes on average per problem. As there were no S5 codes, students legibly presented their ideas, and the limited number of S3 codes implies students typically wrote proofs that followed in a logical order. Across all problems, S8 and S4 codes were the most prominent S codes. There is approximately one S8 code on every two problems. This suggests that students often relied on nonstandard or confusing notation. There is one S4 code on approximately three out of every ten problems, implying the use of extra details, extra steps, or the proof was hard to follow.

Students rarely forgot to conclude their proofs as seen by the small number of U7 codes. Across all problems, approximately seven out of 10 proofs included a crucial step not being sufficiently justified or an important part of the proof not being addressed as suggested by the U4 codes. The U5 codes suggest that about half of the proof attempts included a false statement, incorrect computation or incorrect implication. Finally the average number of U6 codes suggests that approximately three in every 10 proofs did not address some aspect of the problem.

Discussion

1st Level Coding

It appears that students had an understanding of formal proof development and generally made an attempt at mathematical deduction since coded problems were deemed analytical more than 95% of the time. Such results suggest that the course prepared students to think in a mathematically mature manner. It is quite probable that such a high percentage is observed because students saw analytical proofs in class. If a problem was not presented analytically, the class would address this and collaborate on how to construct a more rigorous proof. Only two coursework problems were coded as analytical less than 95% of time; these problems were numbers 4.41a and 5.19. Neither problem was presented, which could imply that it is more challenging for students to formulate rigorous and formal proofs on their after seeing presentation by peers. On the other hand, the remaining three not presented problems were deemed analytical more than 95% of the time. From this, it is logical to conclude that the classroom environment had a positive effect on students and prepared them to attempt formal proofs on their own.

Overall, students seemed to grasp the idea of formal deduction, as they were able to construct analytical proofs for a vast majority of the coursework problems. But were students able to transfer this new understanding of formal mathematical proof into their final examinations? Compared to the coursework, the final problems had a larger percentage (30%) of problems that could not be coded as analytical. This is to be expected since the exam setting is very different than class-time and students were under a time constraint. However 70% of proofs on the final were coded as analytical, which may suggest that the way in which the course was instructed adequately prepared students and gave them the tools to attempt mathematical deduction in future coursework.

Use of Definitions

Students did an impressive job correctly using definitions in their proofs. When considering all problems and groupings, definitions were used correctly more than 95% of the time. There is also little variance (a mere 2%) in percent of incorrect definitions between content areas. These findings are in direct contrast to those of Moore (1994). Two of the seven most common mistakes Moore (1994) found in his research had to do with definitions: D1 (students did not know definitions) and D5 (students did not know how to use definitions to obtain the overall structure of proof). Although Moore (1994) did not provide quantitative data for his conclusions, it seems clear that this study produced results drastically different than Moore's (1994) observations of a traditional lecture-based introductory proofs course.

When comparing the final exam problems to the coursework, students had more than four times as many incorrect definitions in the final. Though this seems like an extreme difference, the reality is that 86% of definitions used on the final were used appropriately. In consideration of the unique circumstances found within the confines of a final exam, one may consider this 86% rate commendable. Recall that students were given brand new definitions on the final and were expected to use them correctly while constrained to three hours and a working environment that did not reflect the learning environment of the course.

In general these students struggled very little with properly using definitions. The classroom environment seems to have helped provide a solid foundation of how to correctly use definitions within a proof, which will be an asset to their future work in mathematical proof.

Use of Assumptions

Moore's (1994) D7 code (the students did not know how to begin proofs), would be translated in the context of this research to high percentages of incorrect assumptions. If students are making false assumptions it is possible they did not have a firm grasp on how to begin formal deduction, since most assumptions are made to start a proof or begin a case within a proof. Since the students in this study had a very low (3.5%) overall percentage of incorrect assumptions, our findings once again seem to contradict Moore's (1994). Furthermore, there is no significant difference (less than one percent) between percent of incorrect assumptions in the average of all course problems and that of all final exam problems.

The only notable variation of assumption use happened among content areas. Approximately 9% of assumptions made in proofs about Relations and Functions were incorrect. These content areas are typically difficult for students at this level and the 9% incorrect may be reasonable in this situation. Even with this slight variation it seems safe to say that students consistently used assumptions correctly. Importantly, there was little variation between the statistics from the coursework and final exam. Thus, it is logical to conclude that the class successfully prepared students to properly use assumptions in future upper division proof-based courses.

Structure (S Codes)

Examining S codes overall, it is apparent that students had few issues with proof structure, language, and notation. A lack of these errors seems to indicate that students were well prepared to organize their ideas and present them in a clear and efficient manner. In contrast, Moore (1994) found that one of the most common difficulties for students was a lack of ability to use and understand mathematical language and notation.

Encouragingly, only 34% of the 473 coded proofs received any S codes. It is logical that seeing problems presented in class with proper notation and structure would make students more likely to do the same on their own. Moreover, since students saw proofs presented by their peers, they also saw mistakes made. Due to the high level of discussion and collaboration in class, these common mistakes were addressed and corrected. This made students more aware of possible mistakes than they perhaps would have been if they had only seen correct proofs completed by their instructor, thus potentially making students more confident and comfortable using correct language and notation when working on their own.

Approximately 2 out of every 5 proofs received an S8 code, indicating confusing notation, while a very low number of other S codes were given. This suggests that there were very few instances of illegibility, illogical order, or the use of extra steps in student proof attempts, which may reflect that the structure of the class taught students how to be concise and argue efficiently.

Unexpectedly, presented proofs had slightly more S codes than not presented proofs. The instructor mentioned that the main focus of in-class discussions was often concepts and not organization. This would explain the higher frequency of S3 and S4 codes, and thus total S codes as well, in presented problems. On the other hand, presented problems had fewer S8 codes (confusing wording or notation), which *is* expected. While seeing correct wording and notation in presented proofs, it seems that students could not always reproduce the notation correctly on proofs just for the portfolio. Also as expected, the proofs for the portfolio have more S8 codes than the homework proofs. It is plausible that because homework was graded while the portfolio was checked for completeness students would try harder to use correct notation on the homework.

When looking at the results by content area, Relations stands out for having a higher average of every S code per problem. The distribution of S codes in the Relations content area was due largely to problem 4.23b, in which students were asked to prove a statement about the composition of relations. Many students seemed to be confused about where to include the statement of the existence of the intermediate variable and this led to an increased number of S4 codes. Also, any time the existence was put in the wrong line the proof was coded with an S3 for illogical order. It is interesting to note that in proofs of the comparable not presented problem, 4.41a, this was not an issue. So it seems that either an error was made in the class presentation that carried over to student proofs or the issue was resolved prior to completing problem 4.41a.

Throughout the finals, S3 codes were also scarce. However, the overall number of S codes is slightly higher on the final exam. Considering the structure of the final and the time constraint, such a slight increase is reassuring. When faced with new material, students continued to have minimal issues with language and notation.

Understanding (U Codes)

The primary focus and purpose of U codes was to assess students' understanding of the necessary steps to prove a statement. On average, students obtained more than one U code per problem and 64% of coded problems contained at least one U code. This may imply students struggled with understanding. However, it seems likely to expect some U errors in an introductory proof class. The most common error regarding implications or steps within students' proofs was U4, which means students were not sufficiently justifying statements. Many student proofs addressed the correct steps necessary to prove a statement but did not explicitly explain how each step was connected. Examining proofs with U4 codes, it seems that students could often recognize the structure of the proof and identify what needed to be addressed in order to construct a rigorous proof but struggled with justification using deductive reasoning.

Although this problem appeared early in the course, problem 2.18 stands out for lack of error with 78% of all proofs having zero U codes. Students were asked to prove an *if and only if* statement about the roots of a quadratic. In a sophomore-level mathematics class, solving problems of this type is already second nature. Since students were confident with what they were proving, it seems they were able to more clearly demonstrate their understanding of the necessary proof technique.

Analyzing the proofs by type and content area provided unique insight into student understanding. Proofs of presented problems had more total U codes than proofs of not presented problems. This may seem contrary to expectation at first glance, however, it is possible when students saw a problem presented they merely copied it down without a clear understanding of why the proof worked (or in some cases did not work). There was also a higher average of total U codes in proofs that were strictly completed for the portfolio. U4 stood out as the main difference. This may be expected from proofs that were not submitted as homework since students knew no grade for understanding and fluidity would be assigned. Proofs in every content area besides Relations had a higher average of U errors than S errors implying the students had more issues understanding how to prove statements than with proper notation.

The average number of U codes per proof for coursework and final exam problems were reassuringly close. As expected, proofs on the final exam had more average U codes, approximately 2 more codes for every 10 proofs. This minor difference in errors could be a result of students' inability to collaborate with peers, the time constraint, or the newness of the mathematical content.

With .06 U7 codes per final exam proof versus .21 per course work proof it should be noted that U7 was the only U code to have fewer average codes per proof on the final. This impressive difference implies that students learned the importance of finishing a proof with a strong conclusion. It is interesting to note that some students who did not prove all problems on the final exam set up the proof with the correct assumptions and conclusion, which may suggest students developed some understanding of the structure required for an analytical proof.

Conclusion

It can be argued the proof abilities demonstrated by students in this class suggest that by the end of the course they were sufficiently prepared for further mathematical study. The high percentages of analytical proofs imply this course provided students with a foundational understanding of formal development. This formal development includes the ability to use definitions and assumptions correctly over 95% of the time. Less than one language and notation error (S code) on average per problem is evidence this IBL class taught students how to convey their thoughts in an efficient and logical manner. Understanding errors (U codes) were more prominent with nearly twice as many U codes as S codes on average. From this it is clear students struggled more with understanding proof techniques than with proof structure. The lack of drastic increase in S and U codes on proofs written during the final exam as compared to coursework problems suggests the course helped develop consistency in student work.

Student success can possibly be attributed to the peer presentations and collaboration that both served as integral parts of the course. Talking to peers and critiquing their arguments seemed to decrease errors, as evidenced by the increased errors on the final. Not only does this collaboration help students succeed during the class, it is also a more realistic imitation of the working environment of mathematicians. Problem solving is often done in the context of open communication between colleagues.

This research has raised some questions for potential further study. First, what should the expectations be from an introductory proofs class? In particular, does the positive outcome related to definitions, assumptions, and S codes outweigh issues with U codes? Second, what role did IBL play in the results of this work? Specifically, would IBL in general foster a contrast to Moore's work or would IBL in other proof based classes yield similar results? Finally, would a lecture-based class, comparable in content and analyzed with the same coding scheme, have findings comparable to this research?

References

Andrew, L. (2009). Creating a Proof Error Evaluation Tool for Use in the Grading of Student-Generated "Proofs." *PRIMUS,* 19(5), 447-462.

Harel, G., & Sowder, L. (1998). Students' Proof Schemes: Results from Exploratory Studies. *CBMS Issues in Mathematics Education,* 7, 234-283.

Moore, R.C. (1994). Making the Transition to Formal Proof. *Educational Studies in Mathematics,* 27, 249-266.

Schinck, Amelie G. (2011). "Why Use IBL? The Road to Present Day Inquiry-Based Learning." Jan. 2011. Web. 28 July 2011. <http://www.inquirybasedlearning.org/AIBL/Why_Use_IBL.html>.

Weber, K., & Mejia-Ramos, J.P. (2009). An Alternative Framework to Evaluate Proof Productions: A Reply to. *The Journal of Mathematical Behavior,* 28, 212-216.

Weber, K. (2010). Mathematics Majors' Perceptions of Conviction, Validity, and Proof. *Mathematical Thinking and Learning,* 12(4), 306-336.

Weber, K. (2005). Problem-solving, Proving, and Learning: The Relationship between. *Journal of Mathematical Behavior,* 24, 351-360.

STUDENTS' WAYS OF THINKING ABOUT ENUMERATIVE COMBINATORICS SOLUTION SETS: THE ODOMETER CATEGORY

Aviva Halani
Arizona State University

This paper aims to address students' ways of thinking about the sets of elements being counted in enumerative combinatorics problems, known as solution sets. Fourteen undergraduates with no formal experience with combinatorics participated in individual task-based interviews in spring 2011. Open coding was used to identify students' ways of thinking about solution sets. One category of ways of thinking which emerged from the data analysis involves holding an item constant and cycling through possible items for the remaining spots in order to generate all elements of the solution set. This category is known as Odometer thinking and two ways of thinking from this category, Standard Odometer and Wacky Odometer, are presented here. The conjectured Generalized Odometer way of thinking, which involves holding an array of items constant, is introduced as an extension of Wacky Odometer thinking.

Key words: ways of thinking, counting problems, combinatorics education, solution set

Introduction and Research Questions

According to Piaget and Inhelder (1975) children's combinatorial reasoning is a fundamental mathematical idea based in additive and multiplicative reasoning. Indeed, as Kavousian (2008) said "without much prior knowledge of mathematics, one can solve many creative, interesting, and challenging combinatorial problems" (p. 2). This indicates that students should be able to solve combinatorial problems by employing their additive and multiplicative reasoning. However, the research indicates that students often struggle to solve combinatorial problems (Batanero, Godino, & Navarro-Pelayo, 1997; Hadar & Hadass, 1981; Lockwood, 2011). In particular, in a study conducted by Batanero et al. (1997), the majority of students both with and without instruction struggled to give the correct answer to combinatorics problems involving one combinatorial operation. Furthermore, there is evidence that post-secondary students must navigate a variety of pitfalls on the road to solving combinatorics problems (Hadar & Hadass, 1981).

In order to address these difficulties, some studies have investigated which formulae students use to respond to particular combinatorial problems (CadwalladerOlsker, Annin, & Engelke, 2011) and student errors (Batanero, et al., 1997; Kavousian, 2008). However, much of the prior research on combinatorics education has focused on students' actions, not on students' reasoning and understanding. It is widely accepted by mathematics educators that the fact that a student can *do* something does not imply that the student *understands*, or that the student is applying coherent reasoning (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002). Thus, it is not enough to examine students' actions as they solve particular combinatorics problems – it is essential to understand their reasoning as well. Further, it will be foundational to understand the stable patterns in reasoning that students apply in a variety of combinatorial situations. These coherent patterns in reasoning are known as *ways of thinking* (Harel, 2008). The research study described here aims to answer the following research question: What are students' ways of thinking about the set of elements being counted in combinatorial problems?

Lockwood (2011) identified two main perspectives of thinking about combinatorial problems: the *process-oriented perspective*, and the *set-oriented perspective*. In the process-oriented perspective, the act of counting equates to completing a procedure which consists of individual stages. The student may or may not associate this procedure with a set of outcomes. In the set-oriented perspective, the act of counting equates to determining the cardinality of the set of objects being counted, known as the *solution set*. Lockwood (2011) claims that being able to coordinate processes and sets is important. She reasons that although thinking in steps or stages is a necessary part of counting, it is sometimes vital to link the process with a set of outcomes. Framed in this language, the research question investigates students' ways of thinking about the solution set of combinatorial problems.

Theoretical Framework

The philosophical perspective underlying this study is that "knowledge is not passively received either through the senses or by way of communication, but it is actively built by the cognizing subject"(Von Glasersfeld, 1995, p. 51). This idea that mathematical knowledge is constructed as the learner engages actively in the tasks is central to this research.

Harel (2008) contends that there are two different categories of mathematical knowledge: ways of understanding and ways of thinking. Humans' reasoning "involves numerous *mental acts* such as interpreting, conjecturing, inferring, proving, explaining, structuring, generalizing, applying, predicting, classifying, searching, and problem solving" (Harel, 2008, p. 3). *Ways of understanding* refer to reasoning applied to a particular mathematical situation – the cognitive products of mental acts carried out by a person (Harel, 2008). For example, consider the mental act of problem solving. The exact solution provided by a student represents a way of understanding since it is the product of the problem solving act.

Ways of thinking, on the other hand, refer to what governs one's ways of understanding – the cognitive characteristics of mental acts – and are always inferred from ways of understanding (Harel, 2008). In the problem solving example above, certain problem solving approaches might become clear as the student progresses through different tasks. These approaches could include "try a simpler example" and "just look for key words." These are ways of thinking since they are characteristics of the students' problem solving acts. Reasoning involved in ways of thinking does not apply to one particular situation, but to a multitude of situations (Harel, 2008). According to Harel (2008), ways of understanding and ways of thinking thus comprise mathematical knowledge.

The ways of thinking described in the Results section are explained in the context of what a student *does* mentally. If a student engages in these mental acts without fully anticipating the final result, we would call the student's solution a way of understanding. On the other hand, if the student can anticipate the result of these mental acts before completing them, we can say that the student is engaging in a way of thinking. It can be difficult to ascertain from a single encounter whether a student is demonstrating a way of understanding or engaging in a way of thinking – it is necessary to delve more deeply and examine students' solutions and approaches to other tasks to determine whether students are engaging in a way of thinking.

Research Methodology

Data for this study comes from a series of individual exploratory teaching interviews (Steffe & Thompson, 2000) conducted at a large southwestern university in the USA. Fourteen students from a second-semester Calculus for Engineers course participated in individual task-based interviews. Each student participated in two 60 minute interviews with the researcher (the author) in a two week period in spring 2011. None of the students had formal experience with combinatorics. The purpose of these interviews was to catalogue students' ways of thinking about the elements of solution sets. Each interview involved the researcher as the teaching agent, one of the students, and a series of tasks. All of the interviews were audio and video-recorded.

In the researcher's pilot study, when presented with a combinatorics question, students would immediately begin to try and solve it. This made it difficult for the researcher to see how the students were envisioning the situation. In addition, it is known that students do not always interpret combinatorial tasks in the same manner that the mathematical community does (Godino, Batanero, & Roa, 2005). As a result, tasks for this study were separated into two parts: a situation and a question (or questions). In general, each task began with the researcher presenting a situation to the student. After thinking about the situation for a few moments, the student would then share what he or she envisioned about the situation. The researcher asked clarifying questions about their responses and then presented the question or questions one at a time. The student was given a few moments to think and then shared ideas with the researcher. The researcher asked clarifying questions to probe the students' actions, ways of understanding or ways of thinking, but only intervened if the student was stuck or once they had solved the problem. Tasks for this study involved the operations of arrangements with and without repetition, permutations, and circle permutations.

There were a few phases of retrospective analysis. Following each interview, the researcher took a few minutes to speak her initial thoughts about the students' ways of thinking aloud while using a pen which records audio and links it to writing to record some notes regarding each interview. She discussed the data with two mathematics education researchers during the study. Content logs including summaries of the video for each task were created for each student following each interview. Relevant portions of the video were transcribed as necessary. At the end of the study, the researcher used open coding (Strauss & Corbin, 1998) to identify and catalogue the ways of thinking in which each student engaged. Finally, the researcher returned to the original data – the audio and video-recorded sessions and the copies of the student work – to confirm her models of student thinking.

Some terminology is necessary for the following sections. In line with English (1993), the term *item* is used to refer to one of the objects involved in the counting process. For example, in a problem involving counting the number of permutations of {A,B,C,D}, A is an item. The term *element* is used to refer to elements of solution sets. In our example of permutations of the set {A,B,C,D}, ACBD is an element of the solution set. In many of the tasks for this study, elements of the solution set can be thought of as having slots. Here, the terms *position* and *spot* refer to a slot. The item in the second position or spot in ACBD is C.

Results

Several different ways of thinking emerged from the data analysis. One category of ways of thinking was present when students partitioned the solution set into disjoint subsets and found the size of each subset. Another category involves students creating a similar problem, determining the size of the solution set of the new problem, and relating this to the size of the solution set of the original problem. A third category, known as *Odometer* thinking, is discussed here.

Consider a 3-digit odometer. First the odometer would hold numbers in hundreds and tens places constant and cycle through digits for units place, thus moving from 000 to 001, 002, and so forth until 009. Then, the digit in the tens place would increase to 1 and the odometer would again cycle through possible digits for the units place, to create 010, 011, through 019. Following this, the digit in the tens place would again increase and the process would repeat until exhaustion of items in the tens place. Thus, all numbers which can be created with a 0 in the hundreds place would have been generated. Following this, the odometer would increase the digit in the hundreds place to 1 and repeat the entire process again: 100, 101, 102, …, 109, 110, …

In a similar manner, this idea of holding something constant can be applied to combinatorial situations and was the motivation for the odometer strategy from English (1991). In that study, young children attempted solve tasks involving dressing toy bears in different colored shirts and pants. The odometer strategy was employed by the children when they dressed the bears in all possible colored pants for a certain shirt before changing the color of the shirt. An extension of this strategy is the Odometer ways of thinking where the main idea is to hold one thing constant and systematically vary the other items to create all possible outcomes with that thing in that placement. The thing being held constant would then change and the process would repeat until all possible outcomes had been generated.

It is important to note that in the Odometer way of thinking students are able to anticipate that this idea of holding an item and systematically varying the other items will generate the set of all possible elements of the solution set. In addition, they must know how to systematically vary the other items. They may do so by recursively applying the same Odometer way of thinking, or by using some other system. Essentially, students engaging in the Odometer ways of thinking will have conceptually constructed a tree diagram (or table in the two-dimensional case) and can anticipate how the branches of the trees will be determined. This does not mean that the students have necessarily constructed a tree diagram as it is entirely possible that they do not have the tools to visually represent their thinking, but rather that they have the ability to organize and generate the elements of the solution set in the same way that a tree diagram might.

It can be difficult to distinguish between whether students are using the odometer strategy, as described by English (1991), or engaging in the Odometer way of thinking. The most important distinction is that students are able to anticipate the result of their mental acts when engaging in a form of Odometer thinking. It is only through probing the students' utterances and actions that the researcher is able to determine if the students have simply stumbled upon a plan of action that is currently fruitful, or if the students are truly engaging in a way of thinking.

In this section, an example of students' preconceptions about Odometer thinking is first provided and then two different versions of Odometer thinking are discussed. Students' solutions to combinatorial tasks driven by the different ways of thinking are presented.

Preconceptions about Odometer Thinking

The Odometer ways of thinking involve holding an item constant and *systematically* varying the other items. In English (1991, 1993), young children often demonstrated the ability to hold an item constant while systematically varying one or two other items. The undergraduate students involved in this study often naturally held an item constant. In many cases, they systematically varied the other items (as discussed in the Standard Odometer and Wacky Odometer sections below). However, when trying to vary several items at once, some students encountered difficulties.

Ricardo was asked to find the number of ways that n people could line up in a row. He was given cards with the letter A-F on them and was not given instruction about whether to begin with specific values of n or to work with a general $n \in \mathbb{N}$. When he was attempting to write out the permutations of {A,B,C,D,E}, he said that he envisioned five different matrices. The first had the letter A in the first column of all of its rows, the second matrix had the letter B in the first column of all of its rows, and so forth. This is evidence that he had the idea to hold an item constant in the first position in each of these matrices.

When the researcher asked him for an example of these matrices, he began with what he called "Matrix A." The cards on the table showed the permutation ABCED and this was the first permutation he put in his matrix shown in Figure 1~~Figure 1~~.

In this matrix, Ricardo always placed A in the first column, indicated to the left of the vertical line. When asked to explain how the other rows were generated, Ricardo stated that he started with B and then moved it "one place, then one place, then one place." As he drew the slanted line in the figure below, he said that it gave a "nice little diagonal (…) of B." In order to create the fifth row, Ricardo went back to the first permutation and began to move the C through the other letters. Notice that he had a little trouble doing this, however, and first wrote B in the last position before changing it to D. He continued this idea of moving one letter through the others until he had 14 rows in his matrix, whereupon he remarked that he had completed Matrix A. However, Ricardo actually missed 10 of the 24 permutations of {A,B,C,D,E} which start with A. This indicates that this way of thinking was not an appropriate way to systematically vary the other items.

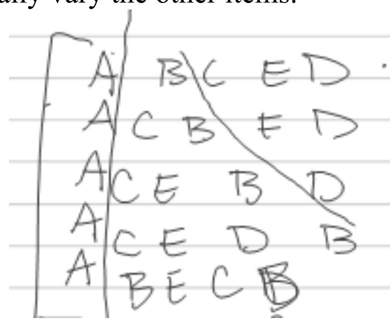


Figure 1: Ricardo's Partial Representation of "Matrix A"

Ricardo was not the only student in this study who attempted to vary items by engaging this way of thinking. Jack described this same way of thinking in the excerpt below and very clearly stated that he only paid attention to the one item he was moving through the others:

> *"It brought me back to like childhood memory of like watching, um, I don't know Disney. An old Disney cartoon where like, they're teaching you something, right? Or, or something. I don't even know how to um, if that's right, but I just remember like visualizing patterns. Maybe like, I visualize each of these cards next to each other, but like one of them moving over [moves the card in the last position to the first position], but it was lit up. That's just what I saw in my head. I don't know why. (...) For some reason, this image of a lit-up letter on a card just kind of. Um, I just saw it um, taking turns [holds one card and moves it through the air] in each spot. (... The other cards) are just kind of moving over. Um, all I can visualize is the lit-up one moving."*

It seems as if Ricardo was attempting to generate each new permutation with A as the first letter by transposing two adjacent items of a permutation already in his matrix. While it is possible to generate all permutations of a set of elements using adjacent transpositions, one must do so recursively. A student attempting to generate all possible permutations in this manner must pay attention to the other items as well. However, Jack's description of this way of thinking and the fact that Ricardo missed 10 of the permutations of 5 letters starting with A indicates that students engaging in this preconception do not attend to the other items in an appropriate manner.

This preconception is included here to show that though students may have the idea to hold an item constant, as Ricardo did, they may not always naturally be able to systematically vary the other items when dealing with a large number of items and positions. However, by understanding students' preconceptions, we may be able to provoke them into developing productive ways of thinking.

Standard Odometer

In the Standard Odometer way of thinking, one would first hold an item constant in a given position and then systematically (and possibly recursively) vary the other items. Following this, the item in the given position is changed and the process repeats until all possible items for the given position are exhausted. The motivation for this way of thinking was the odometer strategy in English (1991).

Example 1. Ben was presented with the Security Code problem below:

- Situation: A security code for a computer involves two letters. It is case insensitive, but the two letters must be different from each other.
- Question: How many possible security codes are there for this computer?

In his solution, Ben anticipated that a security code of the sort AA or BB would not be allowed. He determined the answer to the question to be 26×25. His written work is shown in Figure 3~~Figure 3~~. He explained:

"You have, they have to be different. So if you had the first letter A, it would have to go, you could have A and then B through Z for the next letter. So. And then the same, well the same kind of concept for the next letter was B, you could go A or C through Z..."

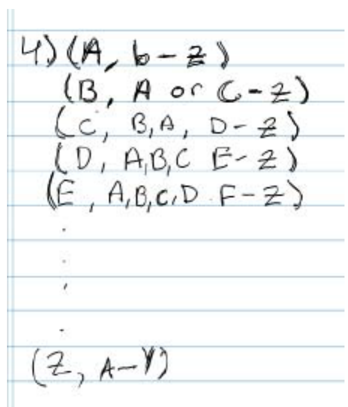


Figure 3: Ben's written work for the Security Codes problem

Ben's written work and explanation seem to be evidence for Standard Odometer thinking. Ben's explanation shows that he first held the A constant as the first letter in the security code. He then cycled through the possibilities for the second letter in the code. Next, he held the "B" constant as the first letter in the security code, and cycled through the possibilities for the second letter in the code. He anticipated that this structure would hold when the letters C – Z were held constant as the first letter in the code, as shown in Figure 1~~Figure 1~~. He recognized that, for each option he held constant as the first letter in the code there were 25 possibilities for the second letter. As a result, since there were 26 possible first letters and 25 possible second letters for each of those first letters, he found the solution to be 26×25.

Example 2. Consider the following Dice problem and Tom's solution below:

- Situation: Two dice are rolled, one red and one white.
- Question: How many possible outcomes are there that are not doubles?

When Tom received the situation of the dice problem and realized we were interested in counting rolls of the dice, he immediately answered, "you have like 36." The researcher asked what he meant, and he responded:

"I can put one here [holds the red die at one] and there are 6 [indicates the 6 sides for the white die]. And then you can change to two [changes the red die to two] and put it with 6 (sides)."

This seems like evidence of the Standard Odometer way of thinking for determining the number of total possible outcomes. He is holding the red die constant at a particular value while varying the values for the white die. He then changes the value on the red die and again varies the white die. It seems clear that he can anticipate that there will be 6 values on the white die for each value on the red die, which is supported by his solution of $6 \times 6 = 36$.

When pressed to explain further, he created the table in Figure 5~~Figure 5~~, writing "1=2" to represent the roll that has a red 1 and a white 2. The researcher then initiated a discussion about whether a red 1 and white 2 was the same outcome as a red 2 and white 1. Tom first believed that this would be true (this explains the crossing out in the figure) but then realized that he was originally correct. The researcher then presented Tom with the actual question. He immediately determined the answer to be 30 and explained that we do not need "1=1", "2=2", "3=3", "4=4", "5=5", or "6=6", so it would be 36-6=30.

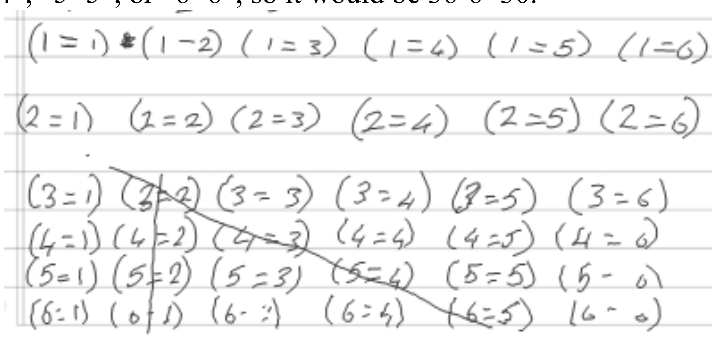


Figure 5: Tom's written work for the Dice problem

Tom's ways of thinking about this task has many similarities to Ben's way of thinking about the Security Codes task, but there are differences. In particular, Tom actually began by constructing a different problem: "How many total possible outcomes are there?" It does seem as if he found the size of this solution set by engaging in Standard Odometer thinking. However, he then had to remove the elements of this solution set which were unnecessary. Thus, Tom was engaging in another way of thinking at the same time.

The Deletion way of thinking belongs to the second category of ways of thinking identified in this study. It involves creating a similar problem, determining the size of the solution set of the new problem, and finding an additive relationship between the size of the new solution set and the size of the solution set of the original problem. Therefore, we can say that Tom engaged in Deletion and Standard Odometer thinking. Indeed, he engaged in Deletion thinking to construct a problem whose solution set was the total possible outcomes, engaged in Standard Odometer thinking to find the solution set of this new question, and found an additive relationship between this new solution set and the solution set he actually wanted by "deleting" the elements he did not want. This is one example of how Standard Odometer thinking might be used in conjunction with another way of thinking.

It is possible that Ben and Tom were not actually engaging in Standard Odometer thinking but instead were engaging in some sort of odometer strategy. However, both students could anticipate the number of items to go in the second position for each of the first positions without actually enumerating them. Furthermore, each student seemed to engage in

Standard Odometer thinking in other tasks as well. Thus it is likely that both students were actually engaging in Standard Odometer thinking.

Tree Diagrams. Neither Ben nor Tom used tree diagrams to visually represent the elements of the solution set. However, tree diagrams can help to show how a student might generate and organize the elements of solution sets. For example, if a student were to engage in Standard Odometer thinking to determine the number of permutations of the letters in the set {A,B,C,D}, he or she could visually represent Standard Odometer thinking about this task using the tree diagram in Figure 7~~Figure 7~~.

In this tree diagram, the student first held A constant in the first slot and then used Standard Odometer thinking recursively to hold items constant in the second positions while varying the items in the third and fourth slots. He or she then changed the item in the first slot and repeated the procedure. Notice that the leaves of the trees are the elements of the solution set and are organized in a lexicographic ordering. A counting process that could be associated with Standard Odometer thinking for this problem would be “There are 4 possibilities for the first slot. For each of those possibilities there are 3 possibilities for the second slot. Then, for each of those possibilities, there are 2 possibilities for the third and fourth slots. Altogether we have $4\times3\times2$ elements of the solution set.” This process counts the number of branches on the tree diagram.

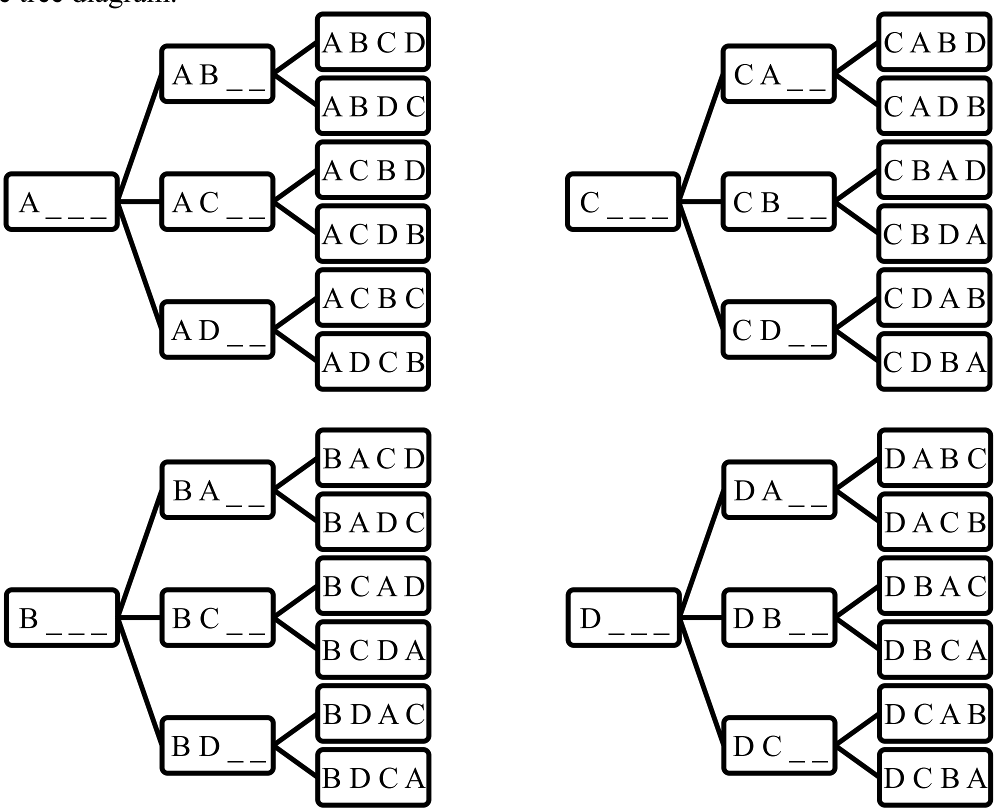


Figure 7: Tree Diagram for Permutations of {A,B,C,D} driven by Standard Odometer Thinking

Wacky Odometer

The Wacky Odometer way of thinking emerged from the data analysis when the researcher realized that sometimes students would hold a single item constant and vary the other items. However, following that, the students would change the position of this item and repeat until all possible elements of the solution set had been generated. This way of thinking is still Odometer thinking since something is being held constant and the other items are being systematically varied, however, it had significant differences from Standard Odometer thinking. It also maintains some similarities to the preconception mentioned above since an item is moving through the others in both; however, in Wacky Odometer thinking, the other items are being systematically varied as well.

In the Wacky Odometer way of thinking, a single item is still being held constant. In contrast to the Standard way of thinking however, the item being held constant each time is not necessarily in the same position. Here, the student would hold one item, say *, constant in a given position and systematically (and possibly recursively) vary items for the other positions. The position of * would then change and the process would repeat.

Example. Jack was asked to find the number of ways *n* people could be lined up in a row. Like Ricardo above, he was given the option of using cards with the letters A-F on them and chose *n*=3 to start. He attempted to find the number of permutations of the items {A,B,C} using the cards with these letters:

> *"So when 'A' is up front, there's two options [moves the cards to create these different permutations]. If A is in the middle [moves the A card to the second position], there's two options. That's two – four. If A is in back [places the A card in the third position], there's two options. Six."*

Jack's explanation indicates that he chose the item A to hold constant in the first position. He then cycled through the possibilities for the items in the other positions, physically doing so in this case. He then changed the position of A, and cycled through the possibilities for the other positions, and repeated a third time. He anticipated that there would be 2 ways to position the remaining items when A was in the second and the third positions.

Jack's way of thinking about this task certainly maintains a similarity to the Standard Odometer thinking in the sense that he held a single item constant in a position and varied the other items before changing something and repeating. However, instead of different items being held constant in the same position, Jack held the same item constant in different positions. Thus, we would say that Jack engaged in Wacky Odometer thinking.

Jack had trouble engaging the same way of thinking for permutations of 4 distinct objects, and instead reverted to engaging in the Standard Odometer way of thinking. He did not use tree diagrams to visually represent his thinking, but his argument was similar to the one described above.

Tree Diagrams. If a student were to engage in Wacky Odometer thinking to determine the number of permutations of {A,B,C,D}, he or she could first hold A constant in the first slot while systematically and recursively varying the remaining items as Jack did in the 3-item case. After changing the position of A, he or she could repeat the procedure for each possible position of A. His or her thinking could be represented using the tree diagram in Figure 9~~Figure 9~~.

In this tree diagram, the student first held A constant in the first slot and then used Wacky Odometer thinking recursively to hold B constant in different positions while varying the items in the two remaining slots. He or she then changed position of A and repeated the procedure. Notice that the leaves of the trees are the elements of the solution set, just as they were with Standard Odometer thinking, but the ordering of the elements is different. A counting process that could be associated with Wacky Odometer thinking for this problem would be "There are 4 possibilities for the position of A. For each of those possibilities there are 3 possibilities for the position of B. Then, for each of those possibilities, there are 2 possibilities for the positions of C and D. Altogether we have $4\times3\times2$ elements of the solution set." This process counts the number of branches on the tree diagram.

The expressions for the number of permutations of {A,B,C,D} that resulted from the tree diagrams generated by engaging in Standard Odometer (Figure 7~~Figure 7~~) and Wacky Odometer (Figure 9~~Figure 9~~) thinking were both $4 \times 3 \times 2$ but the ways of thinking about the task were different, yielding different arguments and different structures on the elements of the solution set.

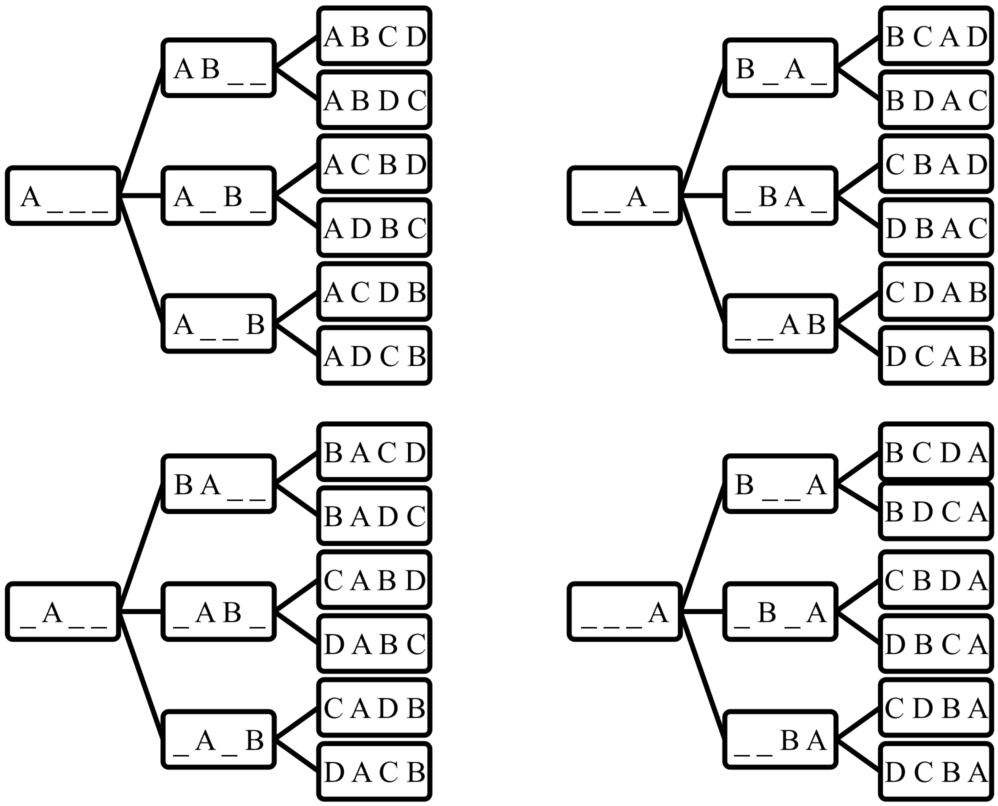


Figure 9: Tree Diagram for Permutations of {A,B,C,D} driven by Wacky Odometer Thinking used Recursively

There are solutions to this task that could be driven by a combination of these two Odometer ways of thinking. For example, a student could engage in Wacky Odometer thinking to place the first letter in each tree, but engage in Standard Odometer thinking to vary the other items. A partial representation of the tree diagram driven by engaging in this combination of ways of thinking is shown below**Error! Reference source not found.**.

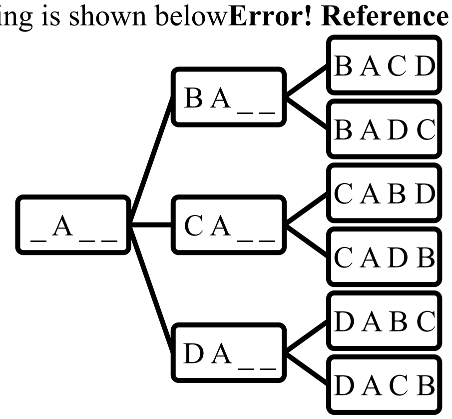


Figure 11: Partial Representation of Tree Diagram for Permutations of {A,B,C,D} driven by Wacky and Standard Odometer used in Conjunction

This student could argue “There are 4 possibilities for the position of A. For each of those possibilities, there are 3 possibilities for the first remaining slot. For each of those, there are 2 possibilities for the two remaining slots.” Again, by thinking about the task in this manner, the student is imposing another structure on the elements of the solution set, and by counting the number of branches in the whole tree diagram, the student determines that there are $4\times3\times2$ total permutations.

Conclusion and Further Discussion

All of the students in this research study naturally had the idea of holding an item constant. Some students struggled to systematically vary the other items while holding one item constant, as Ricardo did when dealing with many items and slots, but most were able to engage in Odometer thinking naturally. In fact, the Odometer ways of thinking were quite prevalent in this research study. Students engaging in Odometer thinking generally tended to the Standard Odometer way of thinking. Jack’s reluctance to engage in the Wacky Odometer way of thinking when the permutation problem became slightly more sophisticated supports this idea.

There are times when Wacky Odometer thinking might be a particularly productive way of thinking about a problem. For example, consider the task “find the number of 3-letter arrangements (repetition not allowed) of the letters *a*, *b*, *c*, *d*, *e*, and *f* which include the letter *d*.” A student attempting to engage in Standard Odometer might engage in Deletion thinking as well to find the number of 3-letter arrangements possible and subtracting the number of 3-letter arrangements which do not include the letter *d*. However, students engaging in Deletion thinking sometimes struggle to construct an appropriate new question to answer. In contrast, a student engaging in Wacky Odometer thinking might vary the position of the item *d*: There are three slots that the *d* could go in, and for each of them, there are 5×5 ways to place the other letters. Thus there are $3\times5\times5$ 3-letter arrangements of the 6 letters which include the letter *d*.

However, there are also tasks for which students tend to naturally choose to engage in Wacky Odometer thinking in a manner which might not be appropriate. Consider the task “find the number of 3-letter arrangements (repetition is allowed) of the letters *a*, *b*, *c*, *d*, *e*, and *f* which include the letter *d*.” A student might attempt to mimic the argument above to say that there are 3 slots that the *d* could go in and for each of those there are 6×6 ways to place the other letters. However, $3\times6\times6$ is larger than the size of the solution set to the problem. By understanding students’ ways of thinking, an instructor might be able to anticipate that students would attempt to find the size of the solution set in this manner and bring students’ attention to why this would over count the elements of the solution set and encourage them to engage in a more productive way of thinking about this solution set.

Since Wacky Odometer can be particularly productive although students may not naturally engage in it, it may be the case that Standard Odometer thinking can be extended to Wacky Odometer thinking. Further, Wacky Odometer thinking may serve as a bridge from Standard Odometer thinking to a more sophisticated way of thinking discussed below.

Generalized Odometer

The Generalized Odometer way of thinking is not rooted in empirical data in this study, but rather is one of the researcher’s own ways of thinking about the solution set of many combinatorics problems. It is an extension of the Wacky Odometer way of thinking in the sense that although things are being held constant, they are not in the same position. However, in contrast to the Wacky Odometer way of thinking, an array of items is being held constant instead of just one item. In this way, it is a more sophisticated way of thinking than either Standard or Wacky Odometer thinking. Consider the following problem and solution driven by the Generalized Odometer way of thinking:

Problem: How many case-insensitive 8-letter passwords are there with exactly 5 E's?

Solution: First, we consider the number of ways to place 5 E's in 8 spots. There are $\binom{8}{5}$ ways to do so. Now consider one of these ways, say E _ E _ E _ E E. Because we can no longer use E's, we only have 25 other item possibilities for each position. Now we can use the Standard Odometer way of thinking (or another way of thinking) to determine the number of ways to fill the remaining positions (25^3). See Figure 12~~Figure 12~~. Note that this was for one possible way of placing the E's. In fact, for *each* way of placing the E's, there are 25^3 ways to fill the remaining positions. Therefore, there are $\binom{8}{5} \cdot 25^3$ total 8-letter passwords with exactly 5 E's.

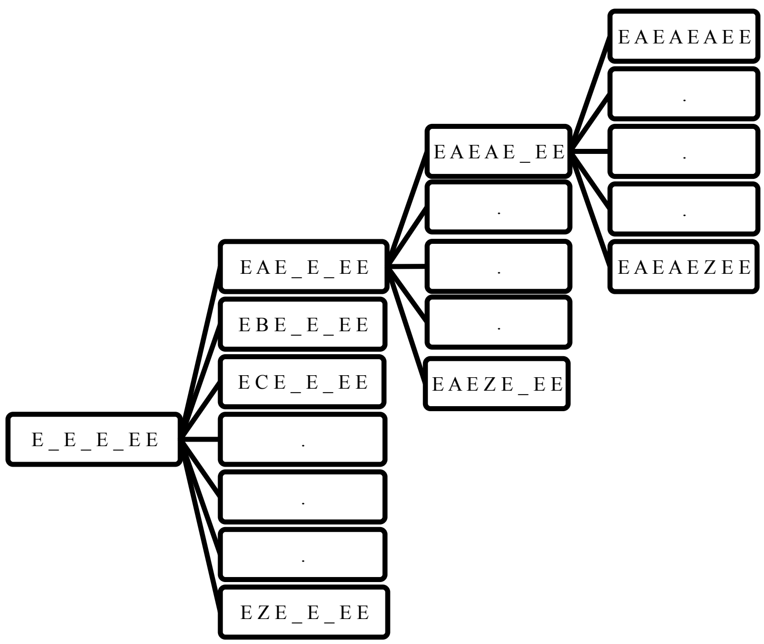


Figure 12: Partial Representation of a Tree Diagram driven by the Generalized Odometer Way of Thinking

Many combinatorics problems can be solved by thinking in stages, but it is important to link those stages with the set of elements as Lockwood (2011) mentioned. In the above solution, the process of choosing where to place the E's and then placing the other letters gives structure to the tree diagram in Figure 12~~Figure 12~~. Thus, the Generalized Odometer way of thinking is a powerful way of thinking that coordinates the process-oriented and the set-oriented perspectives about combinatorics problems identified by Lockwood (2011).

Extending Ways of Thinking about Solution Sets

In many classroom situations, it could be beneficial to provoke students to extend their current ways of thinking about solution sets. For example, if a student such as Ricardo has the idea of holding one item constant in a given position but has trouble systematically varying the other items, it might be productive to encourage him to develop Standard Odometer thinking. If a student engages in the preconception of moving one item through the others without varying those other items, an instructor might wish to encourage him or her to develop Wacky Odometer thinking. Further, since the Wacky Odometer can be thought of as a precursor to Generalized Odometer thinking, it could be fruitful to encourage students to engage in Wacky Odometer thinking before supporting them in deepening that way of thinking to Generalized Odometer thinking. Roh & Halani's (n.d.) instructional provocations might be an effective way to help students build upon their current ways of thinking.

An *instructional provocation* refers to "a discursive move that an instructor may make with the intention to provoke students to further develop their reasoning and understanding by compelling them to re-evaluate their current conceptions or beliefs about a topic" (Roh & Halani, n.d.). Instructional provocations may raise awareness of certain aspects of a topic, highlight inconsistencies or subtle differences in reasoning, raise or resolve cognitive conflicts by presenting new situations, or introduce a new way of thinking about a task for evaluation by students. Four types of instructional provocations suggested by Roh & Halani (n.d.) could be useful to help students build upon their current ways of thinking and develop new ways of thinking: devil's advocate, contrasting prompts, potentially pivotal-bridging examples, and debugging steps.

Devil's Advocate refers to an incorrect or atypical argument or solution provided to students for evaluation. The purpose of this type of provocation is to highlight cognitive conflicts or raise awareness of certain aspects of a topic. After evaluating the argument, the students would either refute the argument or provide justification for portions of the argument. For example, a student might not be aware that it is possible to generate the set of permutations of *n* distinct items by holding one item constant in different places. If this is the case, the instructor might use Devil's Advocate by introducing a solution supposedly written by a former student generating the set of permutations of {A,B,C,D} in the manner of Figure 9~~Figure 9~~, which was driven by Wacky Odometer thinking. The student would then analyze this solution and determine if the reasoning applied in it is logical. If not, the student would refute the argument. If it is logical, the student would justify why this reasoning is appropriate for generating the solution set of permutations of 4 distinct items, and perhaps extend this argument to generating the solution set of permutations of *n* distinct elements. In this way, the instructor is raising awareness of a particular relationship between elements of the solution set through an atypical solution. Further, if the student reflects upon this method and applies it to other situations, he or she may have been provoked into developing Wacky Odometer thinking.

Two statements which sound similar to each other but are not logically equivalent or two statements which are logically equivalent but do not sound similar are called *Contrasting Prompts*. We can extend this concept to have students contrast arguments or solutions driven by different ways of thinking which students might view as legitimate but which are not logically equivalent. For example, a combinatorics instructor might implement Contrasting Prompts by having students contrast two solutions to a combinatorics task, each driven by a different way of thinking. For example, he or she could ask students to determine the number of 3-letter arrangements (repetition is allowed) of the letters *a*, *b*, *c*, *d*, *e*, and *f* which include the letter *d*. If students attempt to engage in Wacky Odometer thinking and determine the answer to be $3\times6\times6$, then the instructor could implement Devil's Advocate by presenting them with an alternate argument written by a supposed former student which is driven by Deletion and Standard Odometer thinking. In this argument, the total number of 3-letter arrangements (with repetition) would first be found by engaging in Standard Odometer thinking and then the number of 3-letter arrangements (with repetition) which did not include the letter *d* could be subtracted for a total of 6^3-5^3. Students might believe both arguments to be reasonable but experience a perturbation when they realize that they evaluate to different numbers. By contrasting the two arguments, their own and the one written by a supposed former student, the students might become more attuned to the elements which are being over-counted in their own argument and why they are being over-counted, thus further developing their reasoning.

According to Zazkis and Chernoff (2008), an example is a "*pivotal-bridging example"* for a student if it pushes the student to re-evaluate their current conception or belief by either raising or resolving cognitive conflicts. The term "pivotal-bridging" comes from the fact that the example then serves as a bridge from the student's initial, naïve conception to a more mathematically appropriate conception. We would say that an instructor is implementing a *Potentially Pivotal-Bridging Example* provocation if he or she introduces an example with the intention of having the student use the example to change their current conception or belief. For example, a student might claim that there are $2n$ permutations of n distinct elements, reasoning based on the number of permutations of 3 distinct elements. The instructor could then suggest a counter-example to the students' conception: the number of permutations of 2 distinct elements. If the student reasons that since a counter-example exists to their claim, they must revise their claim, then the number of permutations of 2 distinct elements would be a pivotal-bridging example for the student. The number of permutations of 2 distinct elements is an example designed to provoke the student to change his or her conception and would therefore be called a "Potentially Pivotal-Bridging Example."

Debugging Steps are delivered as the instructor asks questions or makes statements in order to push the students to test their current conception. The intention of the move is to highlight inconsistencies in student reasoning. For example, a student might claim that there are 2 permutations of the letters A and B: AB and BA, because he could "move" A over to the other side of B create the next permutation. The student might also claim that there are 6 permutations of the letters A B and C: ABC, ACB, BAC, BCA, CAB, CBA, because he could hold one letter constant at the front of his permutation and vary the other two letters and then change which letter is being held constant. If this is the case, the instructor might ask the student if he could apply the "moving one letter over" reasoning to the task of determining the number of permutations of 3 distinct elements. Here, the instructor is adapting the student's way of thinking to a different example. The student will ideally observe how his way of thinking might not apply to more general examples. The instructor's intention is to highlight inconsistencies in the student's reasoning, and so we would say that he or she is implementing Debugging Steps. In this example, the student determined the correct number of permutations in each case; however, the instructor is focusing on the student's reasoning and bringing the student's attention to the inconsistencies.

The key to implementing all of these types of instructional provocations is to first have a model of the students' current ways of thinking. To push them to develop new ways of thinking, it is necessary to have an idea of the desired way of thinking and to understand its constraints as well. Thus, this study which focused on understanding students' ways of thinking about the set of elements being counted and how that thinking expresses itself in their attempts to solve combinatorial problems can be foundational for teaching practice and for future research studies. It can serve to assist teachers in implementing instructional provocations designed to help students develop productive ways of thinking about combinatorics or recognize the constraints of a current way of thinking, and to support curriculum developers in organizing tasks to build upon students' ways of thinking. In addition, this study could provide a framework for analyzing how the ways of thinking are distributed across various mathematical populations. This researcher hopes to conduct further studies to investigate how students develop their ways of thinking about the solution sets as they progress through a variety of combinatorial tasks and the instructor implements provocations designed to encourage particular ways of thinking, including Wacky and Generalized Odometer thinking.

References

Batanero, C., Godino, J. D., & Navarro-Pelayo, V. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*(2), 181-199.

CadwalladerOlsker, T., Annin, S., & Engelke, N. (2011). Counting Problem Strategies of Preservice Secondary Teachers. *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education, 4*, 37-41.

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, 33*(5), 352-378.

English, L. D. (1991). Young children's combinatorics strategies. *Educational Studies in Mathematics, 22*(5), 451-474.

English, L. D. (1993). Childrens' strategies for solving two- and three- dimensional combinatorial problems. *Journal for Research in Mathematics Education 24*(3), 255-273.

Godino, J. D., Batanero, C., & Roa, R. (2005). An onto-semiotic analysis of combiantorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60*(1), 3-36.

Hadar, N., & Hadass, R. (1981). The road to solving a combinatorial problem is strewn with pitfalls. *Educational Studies in Mathematics, 12*, 435-443.

Harel, G. (2008). What is Mathematics? A Pedagogical Answer to a Philosophical Question. In R. B. Gold & R. Simons (Eds.), *Proof and Other Dilemmas: Mathematics and Philosophy*: Mathematical American Association.

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, Portland, OR.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: W.W. Norton.

Roh, K., & Halani, A. (n.d.). Instructional Provocations to Leverage Students' Reasoning and Understanding: A Case of Limit of a Sequence. *Manuscript in preparation*.

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* (pp. 267-307). Hillsdale, NJ: Erlbaum.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (Second ed.). Thousand Oaks, CA: Sage Publications.

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. Washington, D.C.: Falmer Press.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary. *Educational Studies in Mathematics, 68*, 195-208.

INTERCULTURALLY RICH MATHEMATICS PEDAGOGICAL CONTENT KNOWELDGE FOR TEACHER LEADERS

Shandy Hauk
U. Northern Colorado & WestEd

Michelle C. Chamberlin
University of Wyoming

Billy Jackson
St. Xavier University

Nissa Yestness, Kristin King, and Robert Raish,
University of Northern Colorado

We report on our work to build an interculturally aware theory for pedagogical content knowledge (PCK) in the context of teacher leadership. The effort is based on existing and continuing work on developing pre- and in-service teacher classroom PCK and intercultural competence. The RUME session focused on two discussion topics. Discussion Item 1: How do we identify and capture evidence of what might be called "teacher leader pedagogical content knowledge" in interculturally aware ways? Discussion Item 2: What question formats (for written assessments, surveys, interviews) might be productive for eliciting information from teacher leaders about their awareness of and attention to the intercultural aspects of mathematics instruction? ...of mathematics itself?...of teacher leadership?

Key words: pedagogical content knowledge, teacher leader, intercultural competence

Relation of the Work to the Research Literature

Teacher leaders are experienced teachers who take on responsibilities and risks to improve students' educational opportunities while working collaboratively with fellow teachers, administrators, and others (Yow, 2007). Many teacher leaders are mentors to colleagues such as math coaches or facilitators of teacher professional development (Borko, 2004), conduits of communication with administrators, and collaborators on educational policy, research, curriculum product development, and school law (Dozier, 2007; York-Barr & Duke, 2004). Many who identify themselves as teacher leaders report entering leadership positions without any formal training, particularly in adult teaching and learning (Lieberman & Miller, 2007; York-Barr & Duke, 2004). Much of the work of a teacher leader involves negotiating meaning across professional and personal cultural differences.

Several frameworks currently exist for professional contexts that involve understanding, interacting, and communicating with people from various "cultures." In particular, healthcare and international relations groups have generated tools for personal and professional growth based on the theory of intercultural development and communication (Bennett, 1993, 2004; Hammer, 2009). "Culture" can include professional and classroom environments as well as personal or home experience. In this sense, several cultures – sets of values and ways of communicating about them – are involved in doing the work of teacher leadership. A university partnership, the Mathematics Teacher Leadership Center (MathTLC), is investigating the potential for university-based mathematics teacher leadership development that involves a partnering of mathematics disciplinary knowledge growth and leadership learning (this appears to be a relatively unexplored area of collegiate mathematics education research). Members of the MathTLC program include teacher leaders (teachers whose current or near-future job roles include leadership responsibilities), university mathematics and mathematics education professors who are instructors in the program, and graduate student and faculty mathematics education researchers. One goal of the MathTLC project is contributing an interculturally aware theory about pedagogical content knowledge (PCK) in the work of mathematics teacher leaders (TLs). In this work we build on existing efforts related to mathematics classroom teacher PCK (Hill, Ball, & Schilling, 2008; Jackson, Rice, & Noblet, 2011) and intercultural competence development (DeJaeghere & Cao, 2009).

Research Questions

Given the ultimate goal of building a theory for mathematics teacher leader PCK, we started by identifying what might be included under the heading "teacher leader pedagogical content knowledge" (TL-PCK). We have relied on the rich practice-based literature and the available research on teacher leadership, particularly in mathematics and science. The underpinning for the definition of TL-PCK is the nested conception of content and context shown in Figure 1. Mathematics PCK is knowledge for teaching mathematics based in the content-teacher-learner triadic interaction. For mathematics teachers, this triad is represented in Region 1 (math-teacher-student). Teacher PCK about mathematics is in use in Region 1 and PCK is developed by a teacher-as-learner in Region 2 (for example, during a district-offered professional development workshop that uses analysis of the mathematical ideas in a lesson as the base "content" for the workshop). Similarly, teacher leader PCK is knowledge about the "content" that *is* Region 1 and can include knowledge of separate and interlinking processes such as knowledge of mathematics, of students, of teachers, of classroom contexts, as well as integrated concepts such as teachers' PCK, student thinking about mathematics, forms of mathematical discourse, and the nature of socio-mathematical norms. TL-PCK *about* Region 1 is *in use* in Region 2 and may be further *developed* in Region 3. That is, a significant portion of what might be called TL-PCK is associated with knowledge of Region 1 and the implementation/adaptation of it during use in Region 2. Just as many are attending to the role of multicultural awareness and responsiveness for teachers to be effective with students within Region 1 (Gay, 2000; McNeal, 2005), a question for us is the role of intercultural awareness in the packing and unpacking of knowledge of Region 1 as it happens in-the-moment in Region 2 as teacher leaders do their work. This has lead to the driving questions for our current work: How can attention to intercultural competence play a role in the development, assessment, and refinement of TL-PCK? In what ways do self-awareness and awareness of others as cultural support mathematics teacher leadership development?

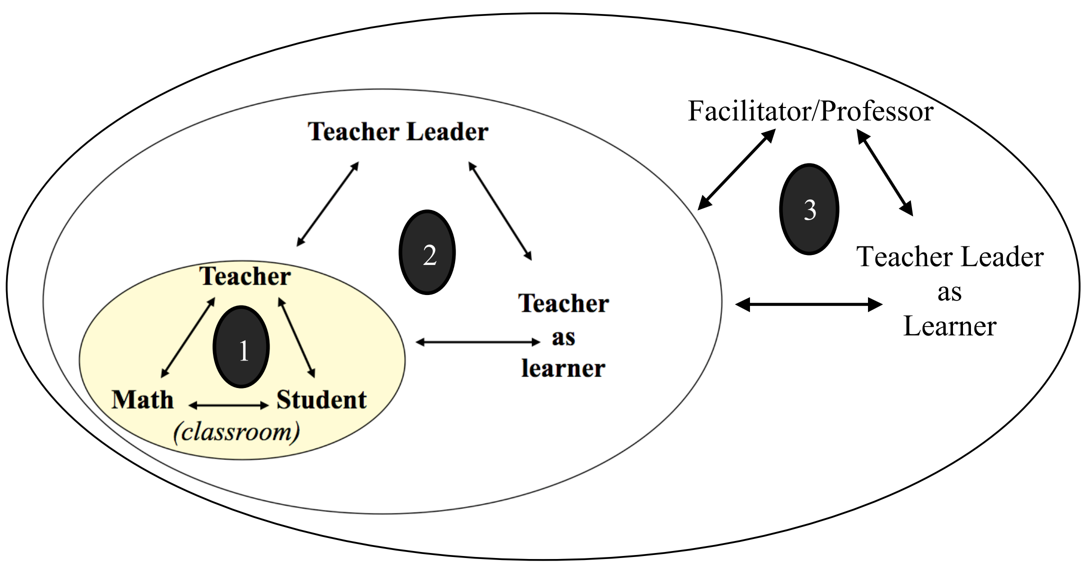


Figure 1. Nested model for teacher leadership

Theoretical Perspective

Our efforts rely on two theories: one for intercultural competence development for mathematics teaching and learning in post-secondary settings and one for PCK. The first is based on the *Developmental Model of Intercultural Sensitivity* (Bennett & Bennett, 2004). A developmental model of orientation towards cultural difference, it includes lower and upper anchor orientations, intermediate orientations, and descriptions of the transitions among the orientations. Associated with the *Developmental Model of Intercultural Sensitivity* in our work is an explicit attention to aspects of discourse based on effective intercultural conflict resolution (Hammer, 2005). See Figure 2 for a visualization we have found useful in describing the stepping places and transitions.

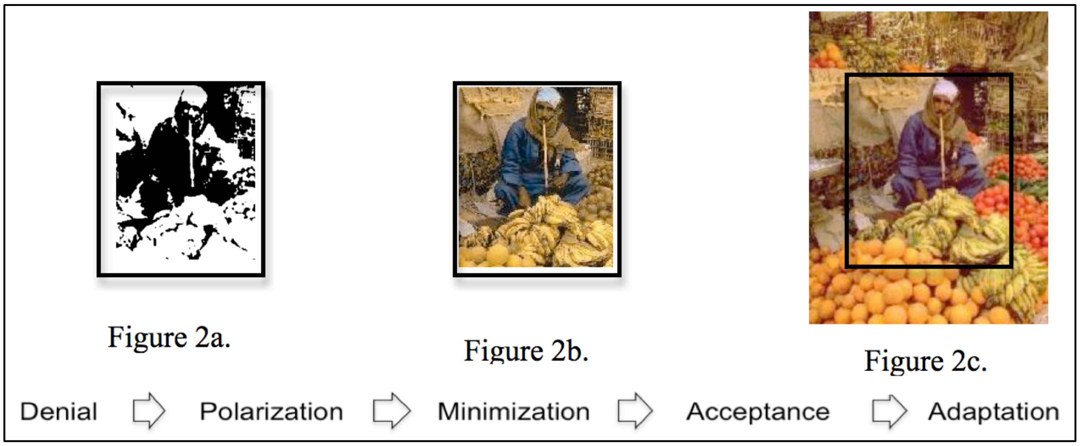


Figure 2. Intercultural competence developmental continuum

The continuum of orientations runs from a monocultural or ethnocentric "denial" of difference based in the assumption "Everybody is like me" to an intercultural and ethnorelative "adaptation" to difference. The development from denial to the "polarization" orientation comes with the recognition of difference, of light and dark in viewing a situation (e.g., Figure 2a). The polarization orientation is driven by the assimilative assumption "Everybody should be like me/my group" and is an orientation that views cultural differences in terms of "us" and "them." A developing tendency to deal with difference by minimizing it by focus on similarities, commonality, and presumed universals (e.g., biological similarities – we all have to eat and sleep; and values – we all know the difference between good and evil and agree on what they are) leads to the minimization orientation. A person in minimization will, however, be blind to deeper recognition and appreciation of difference (e.g., Figure 2b, literally a "colorblind" view). Transition from a minimization orientation to the "acceptance" of difference involves attention to nuance and a growing awareness of oneself as having a culture and belonging to cultures (plural) that differ in both obvious and subtle ways. In the acceptance orientation, people are aware of difference and the importance of relative context, but how to respond and what to respond *in the moment of interaction* is still elusive. The transition from acceptance to "adaptation" involves developing frameworks for perception, and behavior shifting skills, that are responsive to a full spectrum of detail in an intercultural interaction (e.g., the detailed and contextualized view in Figure 2c). Adaptation is an orientation wherein one may shift cultural perspective, without loosing or violating one's authentic self, and adjust communication and behavior in culturally and contextually appropriate ways. There are several ways that knowing one's orientation, or the normative orientation of a group, can inform teacher leader development.

Intercultural theory gives a language for thinking and talking about how we each come to communication. This includes communication across orientations and how we respond to the variety of orientations in a room. The theory also gives a language to develop awareness, to indentify and discuss perspectives about difference and similarity in educational contexts, and for calibrating self-efficacy (e.g., adjust judgments of ability to successfully complete task X to take into account how others involved in task X define "success"). In particular, at the conference we focused on:

RUME Session Discussion Item 1: How do we identify and capture evidence of what might be called "teacher leader pedagogical content knowledge" in interculturally aware ways?

RUME Session Discussion Item 2: What question formats might be productive for eliciting information from teacher leaders about their awareness of/attention to the intercultural aspects of mathematics instruction? ... of mathematics itself?...of teacher leadership? This includes questions for written instruments, interviews, and surveys.

Methods

The work we brought to the conference session is part research and part development. Our continuing research into the nature of professional learning and experience for mathematics teacher leaders and secondary mathematics teachers has included co-development of measures for, and theory around, the knowledge for teaching secondary mathematics as well as the knowledge for mathematics teacher leadership. The focus at the conference was giving a situated view of the theory development for TL-PCK *and* the co-evolving development of measures (written and interview) for TL-PCK.

Our exploration of the intercultural aspect of teacher leadership and the nature of pedagogical content knowledge for teacher leaders is mixed-methods. Quantitatively, we have relied on several existing measures and two project-developed instruments. Qualitatively, our work has included interviews, observations, and examination of documents. For the RUME 2012 session, to contextualize the Discussion Items, we gave an overview of results from several components of the MathTLC research program. The MathTLC program members (teachers, teacher leaders, graduate students, post-docs, and faculty), all completed a 50-item validated and reliable *Intercultural Development Inventory* (see idiinventory.com) that provided intercultural orientation profiles of stakeholder groups. To date, we have completed thematic and categorical coding of teacher leadership program application essays along with initial cognitive interviews and piloting of written assessments of teacher leader pedagogical content knowledge.

Participants in research data gathering for the MathTLC research program have included, to date, 14 teacher leadership program participants (teachers of grades 4-12), 42 master's program students (secondary mathematics teachers of grade 6-12), and 18 university faculty, graduate student mathematics education researchers, and post-doctoral researchers.

Results and Development at RUME 2012 Session

To give a sense of the teacher leader population in the project and a preliminary portrait of TL-PCK and cultural awareness, we summarize analysis of application essays for 14 teacher leaders (the first of four planned cohorts) in Figures 3 and 4. Essays prompts were about (1) ideal classroom, (2) significant experiences prompting a move to leadership, and (3) personal and professional goals.

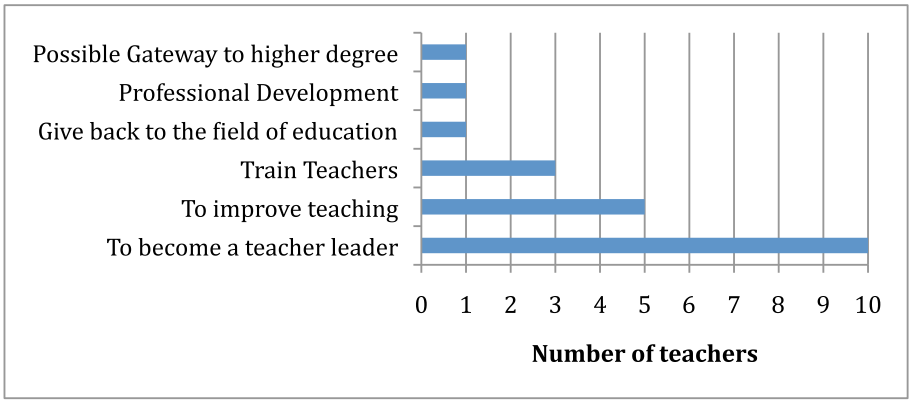


Figure 3. Teacher leader applicant professional learning goals.

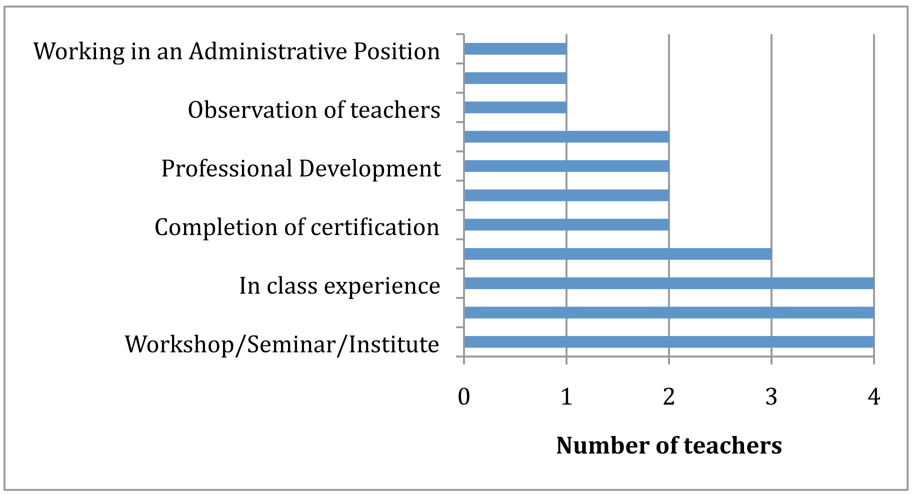


Figure 4. Significant experiences prompting a focus on leadership.

Many TL participants talked about the desire to understand another persons' perceptions: "I hope the program will help me gain a deeper understanding of how other teachers view their teaching of mathematics" and a to "translate my knowledge and skills as a classroom teacher into pedagogical knowledge about adult teachers learning math and learning to teach math to diverse population." Reports on goals included "My hope would be that through my participation in this program I would gain the skills and knowledge to improve my own teaching, better meet the needs of the diverse population of County High School and to influence more classroom teachers to be involved in the school improvement process from the classroom to the national level." For context, we offer also Figure 5, showing the distributions of intercultural orientations of program members along with a reference set of additional stakeholders: secondary mathematics teachers (the "students" of the program's teacher leaders). As a group, the teachers' orientation has been normatively in polarization while the teacher leaders have been largely at the lower end of minimization and university folk largely in minimization.

As part of the research process, we have conducted group profile debriefing sessions with teachers, teacher leaders, and university staff and asked how knowledge of these orientations (for oneself and awareness that they exist for others) might play a part in their professional work. We have also created items used on a written instrument and in interviews with teacher leaders to look at the various aspects of the TL-PCK model shown in Figure 1.

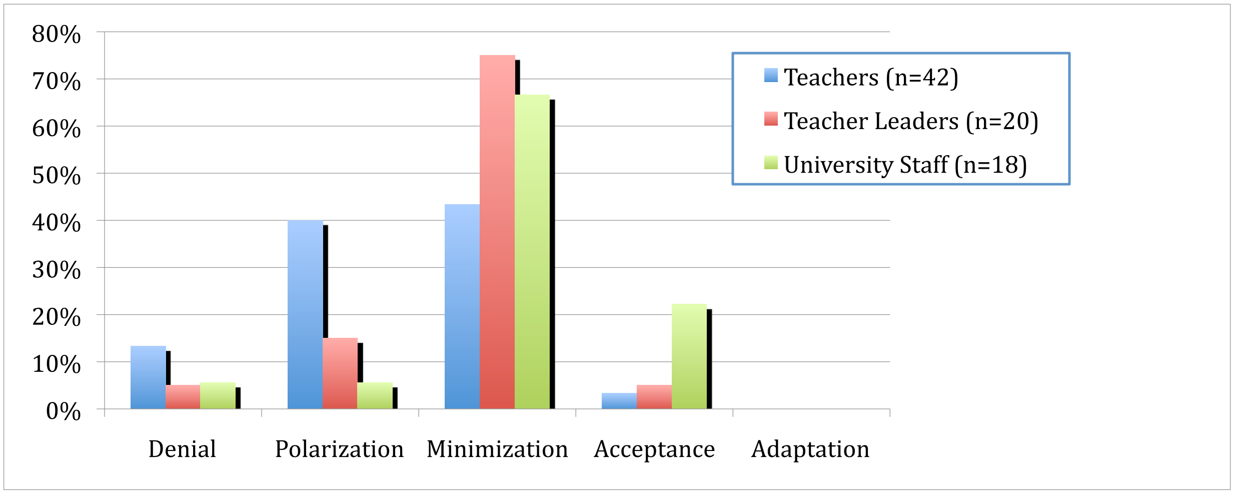


Figure 5. Distribution of intercultural orientations for stakeholder groups

As discussed in the conference session, one of the challenges for the researchers is acknowledging minimization tendencies in developing measurement instruments and attending carefully to nuances in professional cultural differences. Here cognitive interviews with teachers with polarization and acceptance orientations have been most helpful. The noticing of difference by these teachers (both large scale and subtle) has helped researchers acknowledge differences in assumptions about what constitutes mathematical understanding, awareness of others, and the relative importance of these in instructional decision-making. This was foregrounded in the conversation about Example 2. Below, we give several examples along with a summary of the session discussions of the Examples 2 and 3.

Example 1

Part 1. Create a story problem whose solution would require 8th grade students to solve the following for x: $5x - 3 = 12$.

Part 2. What challenges might you expect the students to encounter in doing your story problem?

Part 3. Now think about helping teachers in a PD workshop to build skills in writing story problems. What challenges might you expect 6th to 8th grade teachers to encounter in creating such a story problem?

Part 4. [Given examples of two different teachers problem posing efforts – either on video or in writing] How would you respond to each of the teachers?

Example 2

You are planning a PD workshop on responding to student thinking. The participants are ten 6th grade teachers with whom you work each month. To get an idea of where the group is in making sense of student thinking, you ask teachers to work on a question at the end of the previous workshop (see Figure 6). In looking at teacher answers, you notice that 7 of the 10 teachers answer the question like the example on the left (Figure 7a) and 3 of the 10 answers are similar to the one on the right (Figure 7b).

Question 1: What have you learned about the group of teachers from their answers?

Question 2: How might you use their answers as you make plans for your workshop?

Since the teacher responses offered in Example 2 were distilled from actual teacher answers on a separate assessment that included the item in Figure 6, discussion in the conference session included expressions of concern that the example teacher answer in Figure

Figure 6. Example 2 detail: Multi-part question answered by teachers.

1. Ms. Sepastin reports working with her class on divisibility rules. She told her class that a number is divisible by 4 if and only if the last two digits of the number are divisible by 4; for example, 7,548 is divisible by 4 because 48 is. One of her students asked her why the rule for 4 works. She asked the other students if they could come up with a reason, and several possible reasons were proposed.

1.1. Which of the following statements comes closest to explaining the reason for the divisibility rule for 4?

(A) Four is an even number, and odd numbers are not divisible by even numbers.
(B) Once you subtract the number represented by the last two digits, the number that remains (e.g., 7,500 in the example above) is a multiple of 100, and any multiple of 100 is divisible by 4.
(C) Alternating even numbers are divisible by 4, for example, 24 and 28 but not 26.
(D) It only works when the sum of the last two digits is divisible by 4 (4+8=12, in this example), just like the rule for divisibility by 3.

1.2. Below, for each of the student statements, describe what you consider to be appropriate teacher responses:

(A) That is true. How can you apply that to this situation? What happens when you have even numbers in this situation?
(B) Good thinking. Can you show that works all the time?
(C) That is true. How did you apply that to this situation?
What would happen if you had a really big number and you were trying to test your theory?
(D) That is a good observation about 48. Does it work for 12?

Figure 7a.

None of the students are right, A& C probably because they don't speak English, and B & D probably because they just are talking math words, so I would encourage them to try some examples on their calculators to see if they can find a pattern.

Figure 7b.

7b is mathematically incorrect and that teacher leaders, especially early in a TL experience, might have similar challenges with mathematical content. A generally agreed upon suggestion was that those completing the instrument be asked to do the problem shown in Figure 6 on a separate page before coming to the item as a teacher leader. That is, first ask for activity in Region 1, then ask for activity in Region 2. Such a process is certainly consonant with common mathematics teacher professional development practice – first engage with "content" (do the math) then consider student thinking about the content. Second, given the complexity of attending to intercultural aspects and mathematical PCK of teachers along with mathematics content knowledge of teachers present in Figures 7a and 7b, it was suggested that an interview might be a more productive venue for directing attention as needed than attempting to do it in a written instrument. That is, perhaps teacher leadership participants would complete items like those in Figure 6 (Region 1 engagement) on paper and then revisit the context as a leader (Region 2 engagement) during an interview (possibly revisiting their own work before and/or after engaging with analysis of Figure 7 content).

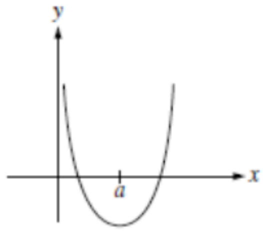
Example 3

Discussion of Example 3 (see Figure 8a, next page), though brief, allowed session attendees to review teacher leader responses to a Region 2-focused item and propose follow-ups to elicit more. Nine of the 10 teacher leaders who responded to the item gave answers like those shown in Figure 8b (next page). These responses seemed to many in the RUME session to be more characteristic of a traditional response one might expect from a teacher to a student (Region 1 activity). The exception was the teacher leader who said: "Explain how you arrived at your answer." How teachers and teacher leaders follow-up with students, when students say or give a correct answer, is an ongoing area of research for the project. This brought up the additional question, during the session, of how teacher PCK activated in Region 2 is repackaged for use in Region 1. For example, besides modeling the behavior, what can facilitators do in Region 3 to scaffold teacher leaders to prompt with "Why" questions to build on teacher correct answers in Region 2 and to have teacher leaders scaffold teachers in asking "Why" questions in Region 1?

Conclusion

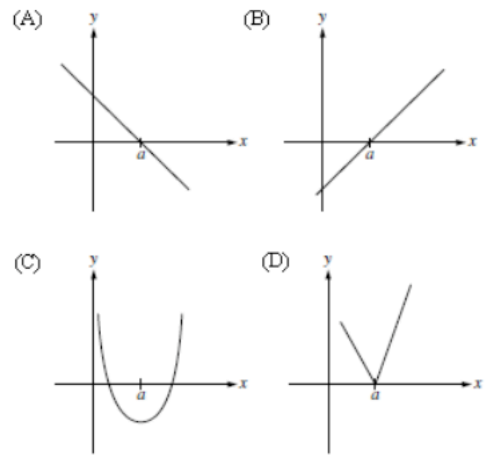
Intercultural orientation is embedded in each content-teacher-learner node and the interaction arrows of the model in Figure 1. That is a great deal of intercultural interaction. How and what a teacher leader notices, how and what a teacher notices, and what a teacher leader does with the noticed things in working with teachers are all connected to self-awareness and other-awareness; linked to the intercultural orientations of all in the professional development classroom. Though beyond the scope of this report, we are also attending to Region 3, the experiences of university teacher-leader educators, whose students are teacher leaders and for whom the "content" is the entirety of Region 2 (including Region 1 as a sub-area).

In thinking about TL-PCK we have relied on the layered model shown in Figure 1, where Region 1 is the "content" in TL-pedagogical "content" knowledge. In our session we talked with the audience about how intercultural aspects of TL-PCK and PCK live in the model. Emergent from the conversation at the conference was the importance, in teacher leadership, of developing the wide area of socio-cultural knowledge needed for teacher leaders to work with administrators, policy makers, and others whose primary work is not itself in Region 1. As we move forward with cognitive interviews we plan to fold in questions about this aspect, which may prove to be orthogonal to the plane in which Figure 1 resides.



Please answer the question below AND address the question that follows:

9.1. The figure above is a graph of a differentiable function. Which of the following could be the graph of the first derivative of the function?



9.2. You pose the question above to teachers in your professional development course. Of them, 68% chose (B), 19% chose (D), and 13% left the item blank. How would you respond to the teachers who chose option B?

Figure 8a.

B is correct because the first derivative graph is the graph of the slopes of the tangent lines. So the slope at a is zero, to the left of a, the slopes of the tangents lines are negative and to the right, the slopes are positive.

You are correct! The derivative represents the slope of the graph which is negative when $x<a$ and positive when $x>a$ and zero when $x=a$.

Figure 8b.

References

Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups* (Tech. Report; NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Bennett, M. J. (2004). Becoming interculturally competent. In J. Wurzel (Ed.), *Towards multiculturalism: A reader in multicultural education* (2nd ed., pp. 62–77). Newton, MA: Intercultural Resource Corporation.

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (2nd ed., pp. 21-71) Yarmouth ME: Intercultural Press.

Bennett, J. M. & Bennett, M. J. (2004). Developing intercultural sensitivity: An integrative approach to global and domestic diversity. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of intercultural training* (pp. 147–165). Thousand Oaks, CA: Sage.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8).

DeJaeghere, J. G., & Cao, Y. (2009). Developing U.S. teachers' intercultural competence: Does professional development matter? *International Journal of Intercultural Relations*, *33*, 437-337.

Dozier, T. K. (2007). Turning good teachers into great leaders. *Educational Leadership*, *65*(1), 54-58.

Hammer, M. R. (2009). Solving problems and resolving conflict using the intercultural conflict style model and inventory. In M. A. Moodian (Ed.), *Contemporary leadership and intercultural competence: Exploring the cross-cultural dynamics within organizations* (pp. 219-232). Thousand Oaks, CA: Sage.

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.

Jackson, B., Rice, L., Noblet., K. (2011). What do we see? Real time assessment of middle and secondary teachers' pedagogical content knowledge. In S. Brown (Ed.) *Proceedings of the 14th annual conference on Research in Undergraduate Mathematics Education.*

Lieberman, A., & Miller, L. (2007). What research says about teacher leadership. In R. H. Ackerman, & S. V. Mackenzie (Eds.), *Uncovering teacher leadership: Essays and voices from the field* (pp. 37-50). Thousand Oaks, CA: Corwin Press.

McNeal, K. (2005). The influence of a multicultural teacher education program on teachers' multicultural practices. *Intercultural Education, 16*(4), 405-419.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research*, *74*(3), 255-316.

Yow, J. (2007). A mathematics teacher leader profile: Attributes and actions to improve mathematics teaching & learning. *Journal of Mathematics Education Leadership*, *9*(2), 45-53

MATHEMATICAL ACTIVITY FOR TEACHING

Estrella Johnson
Portland State University

This work aims to establish a new theoretical construct, mathematical activity for teaching – the mathematical work teachers engage in while teaching. Given this new construct, it is possible to investigate relationships and patterns of interaction between teacher activity and that of their students. By analyzing the classroom video data of mathematicians implementing an inquiry-oriented abstract algebra curriculum I was able to identify four patterns of interaction between mathematical activity for teaching, pedagogical activity, and student mathematical activity. My analysis shows a variety of ways in which teachers' mathematical and pedagogical activity may interact– with some episodes illustrating ways in which these two forms of activity may be somewhat disjoint and other episodes illustrating ways in which these two forms of activity may be tightly integrated.

Key Words: Mathematical Knowledge for Teaching, Mathematical Activity, Teaching Practice, Abstract Algebra

As a way to address the challenges teachers face while implementing reform curriculum, researchers have looked to link Mathematical Knowledge for Teaching (Ball, Thames, & Phelps, 2008) to certain teaching demands. While these studies begin to identify the process by which teachers' knowledge influences their teaching, there remain questions about how teachers' mathematical knowledge directly relates to the mathematical activity of their students. Presumably, it is not enough for teachers to simply *have* the mathematical knowledge that underlies their curriculum. Teachers also need to be able to *use* their mathematical knowledge in a way that supports their students' mathematical activity.

Indeed, in Ball et al.'s (2008) description of their Mathematical Knowledge for Teaching (MKT) framework the authors state that their focus was not limited to knowledge, but also the use of knowledge.

> Despite our expressed intention to focus on knowledge use, our categories may seem static. Ultimately, we are interested in how teachers *reason about and deploy mathematical ideas* in their work. We are interested in *skills, habits, sensibilities, and judgments as well as knowledge*. We want to understand the *mathematical reasoning* that underlies the decisions and moves made in teaching. (Ball et al., 2008, p. 403, italics added)

Here we can see that entwined in this discussion are dual constructs, 1) knowledge and 2) activity – such as skills, habits, pedagogical thinking, and reasoning. With this distinction in mind, questions are raised. For instance: how is teacher activity related to the mathematical activity of their students?

Theoretical Perspective

Throughout the research literature we find examples of specific instructional activities that are believed to be supported or constrained by teachers' MKT. At the undergraduate level, Speer and Wagner (2009) presented a study in which they sought to account for the difficulties a mathematician was facing while trying to provided analytic scaffolding during whole class discussions, where analytic scaffolding is used to "support progress towards the mathematical goals of the discussion" (p. 493). Speer and Wagner identified several components necessary for providing analytic scaffolding, including the ability to recognize and figure out both the ideas expressed by their students and the potential for these ideas to contribute towards the mathematical goals of the lesson. Speer and Wagner went on to state that, "recognizing draws heavily on a teacher's PCK [pedagogical content knowledge], whereas *figuring out requires that a teacher do some mathematical work in the moment* [emphasis added]" (p. 8).

In a somewhat related study, Johnson and Larsen (2012) investigated a mathematician's ability to *interpretively and/or generatively listening* to their students' contributions, where interpretive listening involves a teacher's intent of making sense of student contributions and generative listening reflects a readiness for using student contributions to generate new mathematical understanding or instructional activities (Davis, 1997; Yackel, Stephan, Rasmussen, & Underwood, 2003). While such skills may not necessarily be mathematical in nature, I hypothesize that they may rely on a teacher's ability to engage in certain mathematical activities. For instance, in order to engage in interpretive listening, a mathematician may need to interpret a student's imprecise language, generalize a student's statement into a testable conjecture, or identify counterexamples to a student's claim (see Johnson & Larsen, 2012).

Such examples of teaching practices that are likely to require mathematical work are not limited to research on mathematicians teaching undergraduate mathematics. Studies focused on in-service and pre-service elementary teachers have also identified analyzing student work, interpreting student explanations, and building on student contributions as important instructional activities needed for teaching mathematics (Charalambous 2008, 2010; Hill et al., 2008). It is likely that each of these tasks require teachers to engage in mathematical activity, as such, these tasks are representative of *mathematical activity for teaching* – where mathematical activity for teaching is defined as the mathematical work teachers engage in while teaching.

Here the term *mathematical activity* is used in a manner consistent with Rasmussen et al.'s (2005) *advancing mathematical activity,* where mathematical activity is viewed as "acts of participation in different mathematical practices" (p. 53). With this view, *student mathematical activity* would include: the types of advancing mathematical activity identified by Rasmussen et al. (2005), which includes symbolizing, algorithmatizing, and defining; types of activity that students are likely to engage in as they work to reinvent (Freudenthal, 1991) mathematical concepts, which includes conjecturing, questioning, and generalizing; and types of activity that are associated with proof, which includes evaluating arguments, instantiating concepts, and proof analysis (Larsen & Zandieh, 2007; Selden & Selden, 2003; Weber & Alcock, 2004).

As students engage in such mathematical activity, one would expect that teachers would need to engage in mathematical activity in response. For instance, when faced with a novel proof, a teacher may need to evaluate a student's proof to determine the validity of the argument and possible advantages/disadvantages of this new approach, both in terms of the current task and in terms of their goals for the lesson. Such evaluation may include proof analysis (Larsen & Zandieh, 2007) and identifying connections between the student's proof technique and other mathematical justifications the students would be likely to encounter during the course of the curriculum (see Johnson & Larsen, 2012).

Figure 1 summarizes the kinds of mathematical activity for teaching I have been able to identify by reviewing the research literature.

Mathematical Activity for Teaching

Interpreting students' mathematical reasoning and contributions (Charalambous, 2008; Hill et al., 2008; Johnson & Larsen, 2012: Speer & Wagner, 2009)

- Interpreting students' imprecise language
- Interpreting student explanations
- Making sense of student errors

Analyzing student work, contributions, and proofs (Charalambous, 2008; Johnson & Larsen, 2012)

- Assessing the validity of student strategies
- Generalizing student statements into testable conjectures
- Engaging in proof analysis
- Identifying counterexamples in order to evaluate student conjectures

Figuring out the potential of student ideas to advance the mathematical agenda (Charalambous, 2010; Speer & Wagner, 2008)

- Co-constructing explanations that highlight the meaning underlying procedure
- Building on student contributions to support the students in meaning-making

Figure 1: Mathematical activity for teaching

My initial conjecture is that, in order to support student mathematical activity, teacher activity will be both mathematical and pedagogical. In particular, it seems likely that a teacher's mathematical activity may support students' mathematical activity indirectly in the sense that mathematical activity for teaching would inform their *pedagogical activity*. Such pedagogical activity could include providing counterexamples, stating the formal mathematical version of a student contribution for a class discussion, and exhibiting a proof for the class. However, there are a number of other examples of pedagogical activity that would likely require minimal mathematical activity on the teachers' part. For instance, revoicing, "the reuttering of another person's speech through repetition, expansion, rephrasing and reporting" (Forman, McComrick & Donato, 1998, p. 531) could be the consequence of evaluating a student statement and determining its potential to move the mathematics forward or could be done merely to foster discussion. The former requires mathematical activity for teaching while the later may be carried out without engaging in mathematical activity for teaching.

It is important to note that, just as the teachers' mathematical activity of interest is directly related to pedagogy, the pedagogical activity of interest is directly related to mathematics. As a result, the primary goal of this work is not to differentiate between mathematical activity for teaching and pedagogical activity, but instead to understand the relationships and patterns of interaction between these two forms of teacher activity. Thus, given these three types of activity (mathematical activity for teaching, pedagogical activity, and student mathematical activity), this paper aims to investigate the following questions:

1) What activity (both mathematical and pedagogical) is present in classrooms enacting an inquiry-oriented, abstract algebra curriculum?
2) In what ways do these types of activity interact?

Conceptual Framework and Methodology

To understand ways that instructors engage with an inquiry-oriented, abstract algebra curriculum we have collected data from the classrooms of three mathematicians over the course of two years. During these two years, there have been four implementations of the curriculum.

For each implementation, every class session was videotaped and members of the larger research team took field notes. Two mathematicians, Dr. James and Dr. Bond, have been the focus of the analysis presented here. These two mathematicians were selected because, in addition to video classroom data, both Dr. James and Dr. Bond participated in debriefing interviews aimed at capturing their experiences in class and in using the curriculum materials. While these debriefing sessions were not held with the above research questions in mind, they occasionally offered supporting or contradictory evidence.

Prior to data analysis I conjectured that, in response to student mathematical activity, teachers might engage in mathematical activity for teaching. This mathematical activity for teaching could then inform pedagogical activity, which in turn would influence subsequent student mathematical activity. This hypothesized pattern of interaction (represented in Figure 2) was developed to investigate possible relationships between students' mathematical activity, mathematical activity for teaching, and pedagogical activity, and guided my iterative data analysis process (Lesh & Lehrer, 2000).

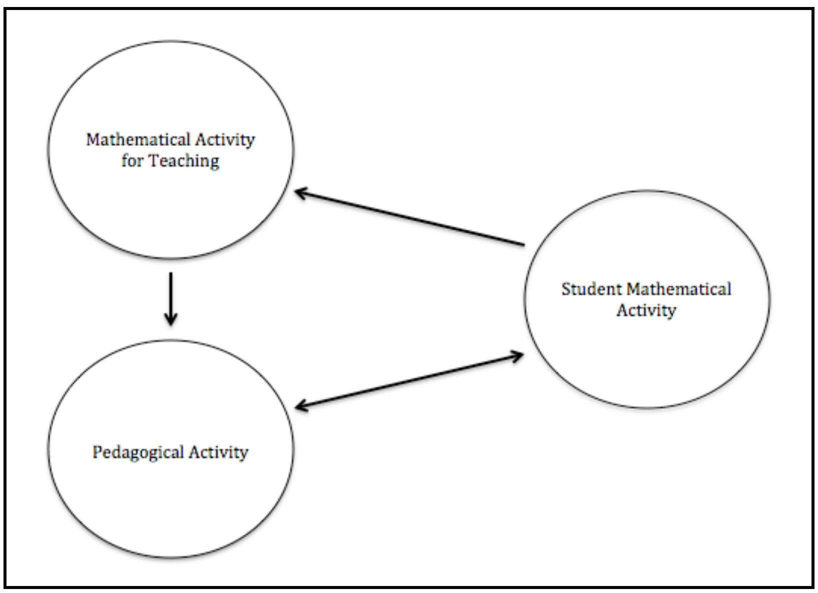


Figure 2: Hypothesized pattern of interaction

Accordingly, I began data analysis by identifying instances in which students would likely engage in mathematical activity by analyzing the curriculum materials. For example, during the unit on groups students are asked to prove some basic theorems related to the order of group elements. Given such a task, I would expect student mathematical activity to include proving. This analysis of the instructor materials served to inform my first round of classroom videotape data analysis. In this round I identified *activity episodes* in which student mathematical activity of interest appeared. These episodes were reanalyzed to see if and how teachers were engaging in either mathematical activity for teaching or pedagogical activity. Finally, I looked for changes in students' mathematical activity following the teacher's activity.

While this process was informed by my hypothesized framework, the framework was open to refinement in light of the research findings. Thus, a wide range of activity episodes were analyzed in an effort to further develop an understanding of the various ways in which these types of classroom activity may interact. As a result, my analysis uncovered several patterns of interaction between teachers' mathematical and pedagogical activity – with some episodes illustrating ways in which these two forms of activity may be somewhat disjoint and other episodes illustrating ways in which these two forms of activity may be tightly integrated.

Results

Here I will present four episodes that illustrate the three components of classroom activity the relationships I am investigating. These four episodes were selected because each illustrates a different pattern of interaction between teacher activity and student mathematical activity.

In the first episode Dr. Bond drew on her mathematical knowledge, but did not engage in substantive mathematical activity for teaching, in order to engage in pedagogical activity. In the second episode Dr. Bond tested a novel student conjecture and in the process became a co-investigator in the mathematics with her students. To do so, Dr. Bond engaged in mathematical activity for teaching. In the third episode Dr. Bond allowed her students' mathematical activity to guide the direction of the lesson. As a result, Dr. Bond's pedagogical activity generated an opportunity for her to engage in mathematical activity for teaching by proving a hidden assumption that her students had uncovered. Finally, in the fourth episode Dr. James engaged simultaneously in both mathematical and pedagogical activity as he navigated a whole class discussion.

Episode 1 – Pedagogical Activity Informed by Mathematical Knowledge

Throughout her course it was common for Dr. Bond to have her students work on proofs individually and then share their ideas within their small groups. Following this group work, Dr. Bond would then initiate a whole class discussion by having a student, or a group of students, share their work with the class. This first episode follows that general pattern.

In this episode the students were asked to prove or disprove that the identity of a group is unique. The students worked individually on this task for about three minutes and then in their small groups for roughly eight minutes. During this time, Dr. Bond circulated the room and briefly interacted with most of the groups. The majority of Dr. Bond's interactions with these groups were evaluative in nature[1], with Dr. Bond quickly reading or listening to student arguments and responding by saying "good". Group work came to a conclusion with Dr. Bond asking Amos to share his proof with the class.

> *Dr. Bond*: All right, let's come together. I think most, if not all the groups, have come up with at least one way to prove this. I have seen two ways as I have gone around the classroom so let's try to share these arguments. Let's see, and um, Amos, can we start with yours? What you just shared with me.

Dr. Bond then used Amos's proof to structure a mini-lecture, where this lecture covered both the proof of this statement and information about how to write and explain an argument for a reader.

In this episode the *students' mathematical activity,* proving, did not prompt Dr. Bond to engage in substantial mathematical activity for teaching. Instead, Dr. Bond was able to draw on her mathematical knowledge in order to *recognize* the student-generated proofs as valid. This

[1] There were two exceptions. With one group of students Dr. Bond discussed what it means to be equal when working with groups that have equivalence classes as elements, as with the symmetry groups. With another group group of students, after determining that they had a valid proof, Dr. Bond asked them to consider if the inverse of a group element was unique.

recognition resulted in Dr. Bond's selection of Amos's proof as the basis of a mini-lecture, which represents *pedagogical activity*. Thus, the activity in this episode can be modeled with the pattern represented in Figure 3.

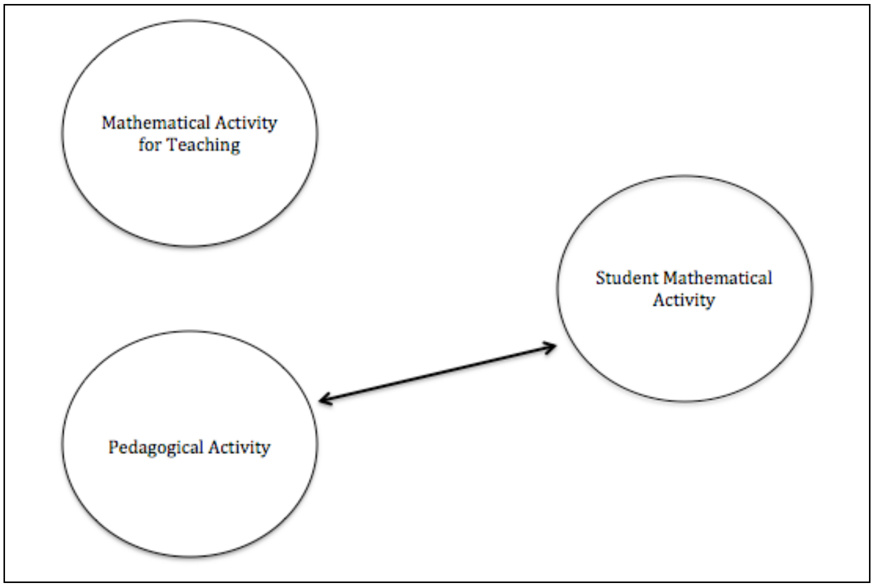


Figure 3. Pattern of Interaction in Episode 1

It is important to note here that Dr. Bond's activity is not void of mathematics. As Speer and Wagner (2009) state, the ability to recognize these arguments as valid "draws heavily on a teacher's PCK [pedagogical content knowledge]" (p. 8). However, nowhere during this episode is it evident that Dr. Bond needed to engage in mathematical work. In particular, there is no evidence that Dr. Bond had to work to *figure out* the students' arguments.

Episode 2 – Classroom Mathematical Activity

As part of the quotient group unit, the students reinvent the notion of coset by considering how they would need to partition a group in order to form a quotient group (Larsen et al., 2009). At this point in Dr. Bond's course the students have been forming quotient groups by breaking dihedral groups into subsets – where those subsets act as group elements. The operation on these elements was set multiplication ($A * B = \{ab|\ a \in A,\ b \in B\}$). The class had already proved that the identity subset needed to be a subgroup and they were asked, given an identity subset, how could they determine what the other subsets needed to be.

While working on this task a couple of groups had noticed that, in each of the quotient groups they had constructed thus far, anytime two elements from the same subset were combined the result was an element of the identity subset.

> *Mark*: I just observed that the one that we worked, in the beginning, the combination of the two elements always gives you an element of the identity set. I have no idea if that actually leads somewhere.

In response to Mark observation, Dr. Bond led the students through a joint exploration of this idea in Z_4.

Dr. Bond began by listing the elements of Z_4, the sets [0], [1], [2], and [3]. She then consider one of these elements, [1], and asked the class "Okay, so what happens… what am I trying to do? So there [referring to the quotient groups constructed from D_8] I, what kinds of things can I try to generalize that argument?" Dr. Bond then leads the class in considering what happens when two elements of [1] are added together.

Dr. Bond: So what happens when I take two elements of the 1 subset and I add them together? What do I get? One plus five gives me? Six, and where does six live?
Class: In [2]
Dr. Bond: Uh, negative three plus five gives me?
Mark: Ah two.
Dr. Bond: In fact I think you find that you always end up in the same place. Now, I think, why are we always ending up in [2]? … It makes sense that, now that we think about it, that one plus five shouldn't get me back to the identity. It also makes sense that R times R^3 should get me back here[2]. Because what do we know about this set? What's the order of it? Who's its inverse?
Class: Itself
Dr. Bond: It's its own inverse. I just had that ah-ha. Because I was thinking about it, I actually thought it might work in Z_n, I hadn't worked it out yet. But it just kind of occurred to me when these all started ending up in [2]. It was like, oh well-duh, because one plus one is two. And that's what's going on there. Those elements were all of order two, so when you multiply them with themselves you're supposed to get the identity back.
Mark: It's just for elements of order two.

We see that during this discussion both Dr. Bond and her students are considering Mark's observation in this new context. Initially Dr. Bond appears to be surprised that the sum of both pairs of elements from [1] was an element of [2]. One explanation for Dr. Bond's surprise is that, up to this point, the quotient group operation had been defined in terms of sets as opposed to representatives (see Larsen et al., 2009). Reconciling this unexpected result appears to have led to Dr. Bond's "ah-ha", as we see with her comment "it just kind of occurred to me when these all started ending up in [2]. It was like, oh well-duh, because one plus one is two".

Therefore, in this episode we see students engaging in conjecturing, which represents *student mathematical activity.* In response to this activity, and Mark's observation, Dr. Bond lead the students through an investigating of the mathematics and in doing so was able to figure out a connection that she did not initially recognize. The testing of this conjecture with her students constitutes *mathematical activity for teaching*. Notice that, while Dr. Bond's mathematical activity was certainly done while teaching, her activity was not substantially pedagogical in nature. Thus, the activity in this episode can be wholly described in terms of student mathematical activity and mathematical activity for teaching, as modeled in Figure 4.

---

[2] Here R and R^3 represent elements in D_8. These two elements made up one of the quotient group subsets that Mark drew on when making his observation.

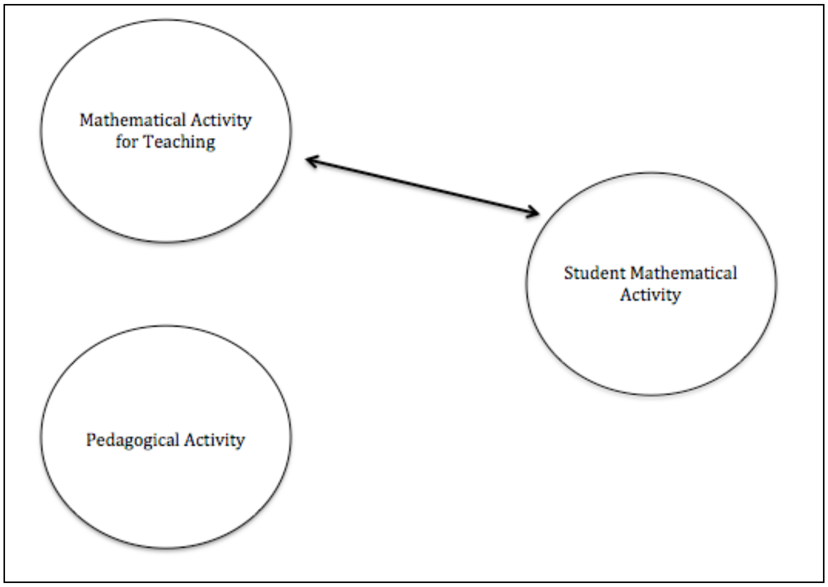


Figure 4. Pattern of Interaction in Episode 2

Episode 3 - Generative Pedagogical Activity

During the deductive phase of the group unit the students were asked to prove that, given group elements a and b, if the order of b is 4 and $ab = b^3a$ then $ab^2 = b^2a$. After a chance to work alone, Dr. Bond asked for volunteers to share their proofs. A student, Tyler, presented a proof by contradiction to the class. In this proof Tyler assumed that $ab^2 \neq b^2a$ and was able to deduce that $ab \neq b^3a$. However, this relied on the fact that if you start with two things that are not equal ($b^3ab \neq b^2a$) and multiply both expressions on the left by the same element, then your resulting expressions are still not equal ($bb^3ab \neq bb^2a$).

Class ended shortly after Tyler's proof, but many students questioned the validity of his proof in a written reflection Dr. Bond collected at the end of class. In these "exit cards" students questioned whether it was valid to assume $bb^3ab \neq bb^2a$ based on the fact that $b^3ab \neq b^2a$. Using these concerns to guide the trajectory of the course, Dr. Bond began the next class by asking the students, "if we take two things that we know aren't equal and we multiply, do we know that they are still not equal"? Initially Dr. Bond stated that this question did not need to be resolved, instead she just wanted to make sure that the students were aware that "this is an important question to ask". Indeed, during the debriefing meeting following this class, Dr. Bond admitted that, " I hadn't decided if it was valid or not … I really hadn't thought it through yet".

However, in the process of raising this question to the class, Dr. Bond gained insight into the justification of the step in question by connecting the student's proof to a previously established result, if $ab = ac$ then $b = c$. She then shared this realization with the class, stating, "my gut at the moment is that … what is, our cancelation property says that if ab equals ac then b equals c, right? And what was the contrapositive to this"? Having made this connection for herself, Dr. Bond was then able to verify the steps of the student's proof with the class.

Given Dr. Bond's debriefing statement, it is clear that this result was not knowledge that she carried with her into class. Instead, Dr. Bond decision to present this hidden assumption to the class provided an opportunity for Dr. Bond to gain insight into its justification. As a result, Dr. Bond was able to spontaneously justify the troublesome step in Tyler's proof. In that way, we see Dr. Bond's pedagogical activity as generative in that it allowed for new opportunities for both mathematical activity for teaching and student mathematical activity. Therefore, in this one episode we see students proving and engaging in proof analysis, which represents *student mathematical* activity; Dr. Bond asking Tyler to share his proof, posing the students' question to the class, and exhibiting a justification, which represents *pedagogical activity;* and Dr. Bond verifying the steps of a novel proof, which represents *mathematical activity for teaching.* This episode is modeled in Figure 5.

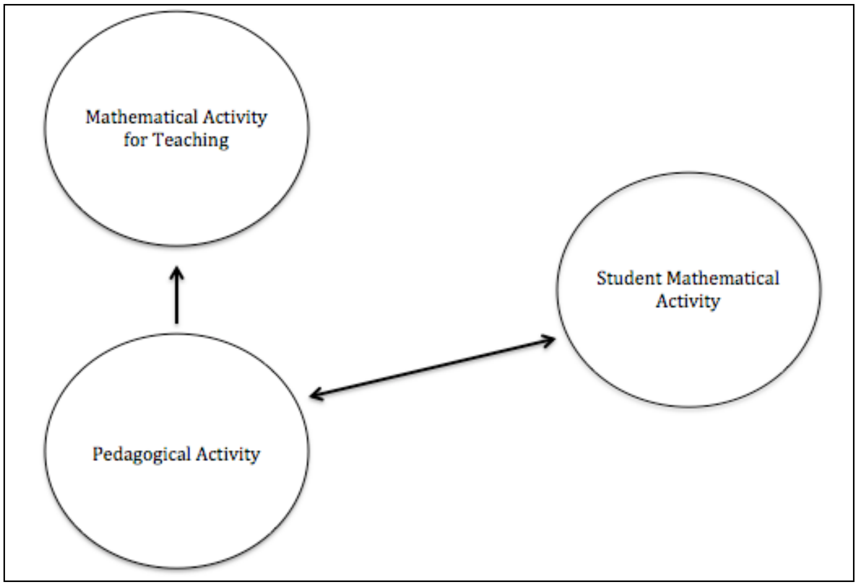


Figure 5. Pattern of Interaction in Episode 4

Episode 4 – Concurrent Mathematical Activity for Teaching and Pedagogical Activity

Dr. James introduced the idea of subgroup by proving that *5Z* under addition is a group in which every element is a member of a larger group, *Z*, under addition. During this discussion students began to question how they know that the identity of *Z*, 0, acts like the identity in *5Z*. For instance, one student asked, "how do you prove it (referring to 0) is the identity of the subset?" Dr. James initially replied by saying, "Well, it's the same operation. So, if something acts like the identity in the whole group it's going to act the appropriate way in the subgroup".

After working through the proof that *5Z* under addition is indeed a subgroup, Dr. James' class then started to conjecture about the minimal list of criteria needed to ensure that a subset of a group is a subgroup. After giving the students time to work in their small groups, Dr. James then asked the groups to report observations or conjectures. One student, Sam, observed that the identity of the group would still act as the identity of the subset. Thus, it would be sufficient to show that the identity of the group is present in the subset, as opposed to proving that the identity of the group satisfies the identity property in the subset. Similarly, Sam observed that the inverse of any element in the group would still be an inverse to that element in the subset. Thus, it is sufficient to show these inverses are present in the subset, as opposed to showing that these elements satisfy the inverse property. At this point, neither of these observations was proved; instead Dr. James was focused on collecting conjectures and observations that the groups would then be given the opportunity to prove.

Shortly after Sam's observation, another student conjectured that it was not necessary to check if the identity was in the subset. Bryan conjectured that, in order to know that a subset of a group is a subgroup, it is enough to check that, for every element in the subset, the inverse of that element is also in the subset, and that the subset is closed under the operation of the group. Bryan argued that, if the subset has the inverse element of every member, then closure will guarantee that the identity element is also a member of the subset, since when you combine any element with its inverse you will get the identity.

> *Dr. James*: So, any element with its inverse, when multiplied together gives you the identity, yeah. That's from closure and from the definition of inverses.
> *Bryan*: And if the identity isn't in your subgroup, then closure would be wrong or if you get a different identity then your group would be wrong.
> *Dr. James*: So that's kind of the general thinking. It's not a full-blown proof. But if you just check closure and you just check the existence of inverses then it seems like maybe that will be enough to get that the identity to be in the subgroup. So, that's an interesting idea that may in fact save us an axiom.

Even though this is not the standard subgroup theorem, it is a common, and valid[3], conjecture that students come up with while working through this curriculum (Johnson & Larsen, 2012). Because this was Dr. James' second time through the curriculum, it is likely that he was expecting this conjecture. However, following this conjecture, the classroom discourse again turned to the identity of the subgroup.

> *Billy*: How do you check inverses without knowing the identity?
> *Eric*: He's saying check the original.
> *Billy*: I'm not sure.
> *Dr. James*: So, talking about checking inverses for… ok, so it's a good point, the identity of the whole group verse the identity of the subgroup.
> *Bryan*: I said that the identity of the subgroup has to be the identity of the group, because if they differ then you have two identities in the main group[4].
> *Dr. James*: Ok, this group was harping on these two things as well. This is probably something we should talk about, clear the air on, and then forevermore be happy about. If you have a subgroup, is it or isn't it, let's actually detour for just a moment. Let's take a few minutes and try to write, everyone, some sort of a proof or something, that if you have a subgroup inside of a group, true or false, the identity of the subgroup has to equal the identity of the group. And, if you can crank that out then consider the second question. Does the inverse of an element in H then have to be equal to the inverse of the same element in G?

In this exchange we see Billy questioning another student's justification for this conjecture – that if the subset is closed and contains the inverse of each element then that the subset will contain an identity element. Billy's challenge is based on the fact that inverses are defined in relation to an identity element, so without knowing that the subset contains an identity element how could you know that the subset contains the inverse of each element. To this challenge Bryan responds with additional justification to provide backing (Toulmin, 1969). Thus, in this episode we can see students engaged in conjecturing, justifying, and evaluating arguments - all

[3] Technically, of course, one must also assume the subset is non-empty

[4] This is incorrect since, by itself, the statement that an element is the identity of a subgroup does not include a claim that it acts as the identity for elements outside the subgroup.

of which constitute *student mathematical activity*.

Clearly evident in this episode is Dr. James' *pedagogical activity.* In response to Billy's challenge, Dr. James presented the class with two questions to be proved or disproved – 1) Does the identity of a subgroup have to equal the identity of the group that contains it, and 2) Does the inverse of an element in a subgroup have to be equal to that element's inverse in the group? The presentation of these two questions to the class represents *pedagogical activity,* which resulted in a new *student mathematical activity* – proving.

Less evident in this episode is a clear expression of *mathematical activity for teaching.* One may argue that Dr. James was simply revoicing either Sam's observation or Bryan's response to Billy's challenge. Let us consider each of these possibilities in turn. Notice that Sam's observation states that 1) if the identity of the group is in the subset then it will act as the identity of the subset, and 2) for any element in the subset, that element's inverse from the group, if in the subset, will act as an inverse. The tasks posed by Dr. James go a step further, by proposing that the only element that could act as the identity of a subset is the identity of the group and that the only possible inverse for an element in a subset is that element's inverse in the group. Thus, the assignment of these two tasks reflects a sophisticated understanding of how these ideas are related but not equivalent and cannot be described as a revoicing of Sam's observation.

Further, in considering Bryan's response, we can see in the transcript that Bryan's response to Billy's challenge only focused on the identity. "I said that the identity of the subgroup has to be the identity of the group, because if they're different then you have two elements that are the identity of the main group". However, Dr. James gave the students two tasks, one to prove or disprove Bryan's claim and another to prove or disprove a related statement about the inverse elements in a subgroup. The assignment of this second task reflects an understanding of how the pieces of the students' mathematical arguments can come together to prove the subgroup conjecture. Thus, Dr. James' activity cannot be interpreted as purely pedagogical.

Therefore, the assignment of these two tasks to the class is both representative of *pedagogical activity* and *mathematical activity.* This stands in contrast to the episodes from Dr. Bond's class and suggests a different way in which mathematical activity for teaching and pedagogical activity may interact with each other. In this episode Dr. James engaged in *concurrent mathematical and pedagogical activity* in response to the student mathematical activity in this lesson, as represented in Figure 6.

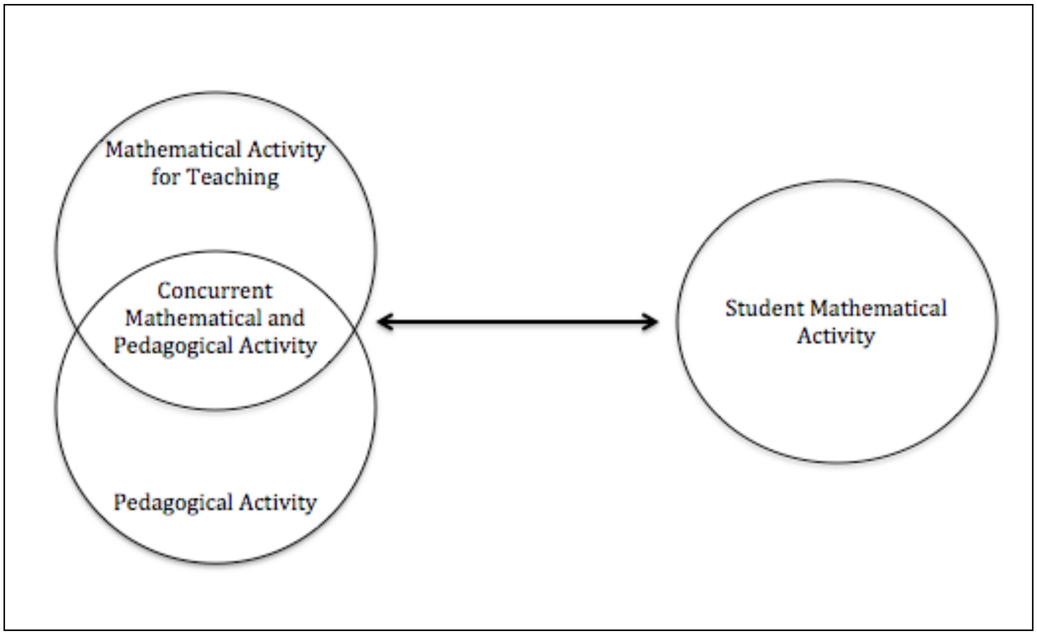


Figure 6. Pattern of Interaction in Episode 4

Discussion

The four episodes presented were selected to illustrate a variety of patterns of interaction between the different types of classroom activity. Specifically, these episodes were selected to highlight ways in which the two types of teacher activity, mathematical activity for teaching and pedagogical activity, may interact with each other as teachers engage with and support student mathematical activity. As evident in these episodes, teachers can engage in pedagogical activity and mathematical activity for teaching either independently or concurrently.

In both of the first two episodes we see Dr. Bond engaging primarily in just one of these types of teacher activity. In the first episode Dr. Bond drew on her mathematical knowledge in order to select a student's proof to share with the class. The selection of this proof was more a matter of recognizing a valid (and routine) argument than a matter of figuring out the student's argument. Thus, Dr. Bond did not engage in substantive mathematical activity for teaching. Instead, Dr. Bond's activity in this episode is representative of pedagogical activity. This is contrasted with the second episode in which Dr. Bond explored a novel student observation with her class. In this episode, Dr. Bond became a co-investigator in the mathematics with her students and made a mathematical connection that was not initially clear to her. In doing so, Dr. Bond primarily engaged in mathematical activity for teaching.

While the first two episodes represent primarily one type of activity or the other, the last two episodes begin to identity ways in which these types of teacher activities can interact with each other. In the third episode there was a clear delineation between Dr. Bond's pedagogical and mathematical activity. Indeed, Dr. Bond's pedagogical activity generated an opportunity for her to engage in mathematical activity for teaching – the proving a hidden assumption that her students had uncovered. Finally, in the fourth episode we see Dr. James pose two tasks to his students. While the posing of these tasks is certainly pedagogical, the development of these tasks is evidence of Dr. James' mathematical activity for teaching. Thus, in this episode it is impossible to tease apart Dr. James' pedagogical activity from his mathematical activity for teaching.

This work offers a first step in defining, identifying, and relating different types of classroom activity. While teacher activity has been discussed in the literature, mathematical activity for teaching has not been defined or differentiated from other types of teaching tasks. Thus, the development of the mathematical activity for teaching construct provides researchers a lens to study teacher activity. However, investigation into both the types of classroom activity and the relationships between these types of classroom activity has only just begun. Most notably, the analysis presented here does not emphasize the role of student mathematical activity or the impact of teacher activity on student mathematical activity.

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.

Charalambous, C. Y. (2008). Preservice teachers' mathematical knowledge for teaching and their performance in selected teaching practices: Exploring a complex relationship Unpublished Dissertation. University of Michigan

Charalambous, C. Y. (2010). Mathematical knowledge for teaching and task unfolding: An exploratory study. *The Elementary School Journal, 110*(3), 247-278.

Davis, B. (1997). Listening for differences: An evolving conception of mathematics teaching *Journal for Research in Mathematics Education 28*(3), 355-376.

Forman, E. A., McCormick, D. E., & Donato, R. (1998). Learning what counts as a mathematical explanation. *Linguistics and Education*, 9(4), 313–339.

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

Hill, H., Blunk, M., Charalambous, C., Lewis, J., Phelps, G., Sleep, L., et al. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study *Cognition and Instruction, 26*, 430-511.

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., Johnson, E., Rutherofrd, F., & Bartlo, J. (2009). *A Local Instructional Thoery for the Guided Reinvention of the Quotient Group Concept* Paper presented at the Twelfth Special Interest Group for the Mathematical Association of America on Research in Undergraduate Mathematics Education Raleigh, North Carolina

Larsen, S., & Zandieh, M. (2007). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67*(3), 205-216.

Lesh, R., & Lehrer, R. (2000). Iterative Refinement Cycles for Videotape Analysis of Conceptual Change. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 992). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Selden, A., & Selden, J. (2008). Overcoming students' difficulties in learning to understand and construct proofs In M. Carlson & C. Rasmusen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* MAA.

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.

Toulmin, S. (1969). *The Uses of Argument* Cambridge, England Cambridge University Press.
Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*, 209-234.
Yackel, E., Stephan, M., Rasmusen, C., & Underwood, D. (2003). Didactising: Continuing the work of Leen Streefland. *Educational Studies in Mathematics 54*, 101-126.

USING DESIGN RESEARCH TO DEVELOP AN ACTIVITY TO INTRODUCE LIMIT OF A SEQUENCE: "GETTING BACK TO SESAME STREET"

Karen Allen Keene
North Carolina State University

William Hall
North Carolina State University

Alina Duca
North Carolina State University

This report is based on work completed within an ongoing project to develop a calculus course which serves as the foundation for the mathematical education of STEM-focused elementary teachers at a large southeastern university. In the process of designing and implementing the course materials, several research-based activities have been developed, tested and refined. In this paper we discuss how we used a design research approach to create and implement an activity that introduces the concept of limit of a sequence using popular characters from Sesame Street. We report on the first two design cycles in the ongoing design of this activity and discuss the modifications made in both the broad learning goals and the activity drafts.

Keywords: Calculus, Limits, Preservice Teachers, Design Research, Sequence

Introduction

The recent emphasis on teachers' mathematical knowledge for teaching is a response to the issue of teachers' needing mathematical content knowledge (Ball, Hill & Bass, 2005; Fennema & Franke, 1992; Mewbern, 2001; Papert, 1971; Thompson, 1992). This knowledge includes not only what is considered common content knowledge, but also specialized content knowledge, i.e., knowledge of mathematics that is specific to the needs of teachers (Ball, Thames & Phelps, 2008). Furthermore, recent research has begun to show that elementary teachers who demonstrate specialized content knowledge do positively impact student achievement (Hill, Rowan & Ball, 2005). In fact, the National Mathematics Advisory Panel (2008) noted,

> Teachers must know in detail and from a more advanced perspective the mathematical content they are responsible for teaching and the connections of that content to other important mathematics, both prior to and beyond the level they are assigned to teach (p. xxi).

Our research addresses both the "more advanced perspective" as well as the connections to other relevant mathematics mentioned by the NMP. The CELTIC project (Calculus for ELementary Teachers: An Innovative Context) is a partnership between three departments, Mathematics Education, Mathematics, and Elementary Education with the common goal of creating a calculus course for preservice elementary teachers. Therefore, striking a balance between the rigorous mathematics of an advanced perspective while taking advantage of lessons from education research is a primary objective of our research. Unfortunately, elementary teachers often lack such an advanced perspective. In their study of the mathematical behavior of preservice elementary teachers, Seaman & Szydlik (2007) found that "teachers display a set of values and avenues for learning mathematics that is so different from that of the mathematical community and so impoverished, that their attempts to create fundamental mathematical understandings often meet with little success" (p. 179). However, as important as rigorous mathematical practice is for preservice teachers to participate in, strict mathematical formalities in calculus for preservice teachers may not be desirable. Wu (2006) calls mathematics education, "mathematical engineering, in the sense that it is the customization of basic mathematical principles to meet the needs of teachers and students" (p. 3) and stresses the importance of mathematicians partnering with educators in order to build appropriate mathematics for K-12 classrooms. This necessity of considering both is critical to our work with preservice elementary teachers.

Why Calculus?

Calculus is the standard mathematics content course required for students at the university level who intend to study in the fields of science, engineering, and mathematics. If one looks across many universities, there are a myriad of calculus courses tailored to meet the needs of those in the STEM disciplines. There are also calculus courses for students entering business, life sciences, or liberal arts. With the increased demand from government and society to improve learning in the STEM disciplines, as well as increase the numbers of people entering STEM fields of work (PCAST, 2012), one elementary education program at a large university decided to address this critical need by requiring calculus for future elementary teachers. The program has been structured to address the goal of educating *STEM-focused preservice elementary teachers*. Developing a calculus course that meets the needs of future elementary teachers, but also provides the students with a calculus base became a priority.

What does this kind of calculus course offer its students? First of all, with the development and adoption of the Common Core State Standards (2010), teachers are being asked to implement instruction that supports the eight "Standards for Mathematical Practices" at all levels, from Kindergarten through 12th grade. These practices invite prospective new teachers to learn mathematics, and then ultimately teach it from a new deeper perspective. The mathematics they will teach is as much about practice as it is about the content. Included in these practices are such items as attention to precision, reasoning about mathematics, modeling with mathematics, and strategically using appropriate tools. While students learn mathematics that is new and challenging, they will experience participation in a class that focuses particularly on these practices. Secondly, students who take and successfully complete a calculus course will ultimately have a deeper understanding of a subject that is standard for high performing STEM future professionals.

Mathematics as a discipline is much more than a list of procedures and memorizing from flashcards. It has structure and is interrelated at many levels. Seeing the calculus and then connecting it to the concepts of number and operations, rational number, etc. provides these students with an experience that will support their understanding of all those concepts. These components of the course specifically tailored for students who prepare to become STEM-focused elementary teachers will increase their confidence in mathematics learning, and thus influence the way they choose to teach mathematics at the elementary level.

We as authors present results from this ongoing project to develop, test, and refine materials for this innovative calculus course for preservice elementary teachers. In the process of designing and implementing the new materials, several research-based activities were developed, tested and refined. Here we specifically focus on the "Getting Back to Sesame Street" activity created to introduce the concept of limit of a sequence and the design research used to construct, test, and refine the activity.

Theoretical Perspective of Learning

We choose to think about learning using the emergent perspective (Cobb, Yackel, & Wood, 1996). Student learning can be studied from both an individual perspective and a collective perspective. Students construct their own meaning for mathematical concepts; sometimes this construction is mitigated primarily through internalizing social behavior, and sometimes these interactions are the milieu in which individuals build meaning (Tudge & Rogoff, 1989). Alternatively, individuals contribute to and learn with others. Participation in a classroom is learning according to this view. Because of this perspective, we believe it is appropriate to study the individuals, the small groups, and classroom activity as a whole to best understand student learning and scaffold the learning in appropriate ways.

What is Design Research?

Design research is a relatively recent methodology that is being adopted by the educational research community. This section provides a brief literature review of design research in mathematics education and in general. The term design research in education was first used by Brown (1992) and Collins (Collins, Joseph & Bielaczyc, 2004). Brown used the term "designer of educational environments" to describe herself and then used that notion to conduct experimental research in classrooms and make improvements based on results. This was a revolutionary approach twenty years ago, but currently has become an accepted and valuable form of research. Those practicing design research build instructional materials in a particular domain and conjectures learning theories about that domain (Cobb, Confrey, diSessa, Lehrer, & Schaubel, 2006). As the experimental research unfolds researchers continuously refine their learning conjectures and materials to assure that goals and materials agree in the final product.

Stylianides and Stylianides (2009) defined an instructional sequence as a series of activities and associated instructor actions for implementation that are developed with design research. During work with elementary preservice teachers to develop their understanding of concepts of limit, we focus on three main practices within design research: incorporating existing research on learning in the creation of an instructional sequence; evaluating the effectiveness of the instructional sequence to support the learning process from the collection and analysis of data from the implementation; going through several cycles of design, implementation, analysis, and refinement to better understand and support learning. Each of these three features will be described in detail in the next section.

As is often the case in design research, the primary researchers in this project also serve as instructors in the calculus course. As teachers, we have a strong commitment to the project; however the dual-role of teacher/researcher may limit our ability to be objective in analyzing the materials. Even with these concerns, however, we feel that the research has been successful and allowed for the emergence of several activities that will ground the entire set of calculus materials.

Design Cycle One

The first design cycle consists of five phases. In the first phase we reviewed the literature concerning the teaching and learning of limits (both limits of functions and sequences). This review of relevant literature established a firm foundation for the second phase, creating a list of broad learning goals for the activity under development. In the third phase we drafted the activity to address each of our learning goals while attending to the recommendations from the literature and our theoretical perspective. While phases two and three are presented as distinct processes, they were actually somewhat simultaneous. We did have some explicit goals before beginning development, while others only came about after we analyzed the activity. Implicit goals were then identified and included in the list of broad learning goals. In the fourth phase, the materials were piloted with volunteer students. The final phase consisted of revising the broad learning goals and activity.

Phase one: Background research

Limit is an often difficult, yet foundational concept in calculus. Underlying both differential and integral calculus, limits can act as the conceptual glue for student learning in advanced mathematics. Unfortunately, limits have proven to be quite difficult for students to learn. Not only do students rarely learn the formal definitions of limits of sequences and functions, but even informal notions of limit tend to be incomplete or inaccurate (Davis & Vinner, 1986; Mamona-Downs, 2001; Oehrtman, 2008; Roh, 2008; Tall & Vinner, 1981; Williams, 1991). While difficult, some educators have found success in limit instruction. Roh (2008) developed activities utilizing "epsilon strips" to help students connect their dynamic views of limit with the formal definition, while Oehrtman (2008), via his approximation framework, has shown how students have intuitive notions about limits of sequences that align with the formal definition of such limits. Here we present a review of such literature as it informed the construction of our limit activity. We note that the research we review may be on function or sequence limit and we are careful not to interpret student understanding of limit of function as understanding of limit of sequence. However, we believe there is something to be learned from both bodies of research.

Student Conceptions of Limit of a Sequence

Researchers have outlined a number of ways students come to understand limits of sequences. We believe it is important to consider these myriad conceptions of limit as well as the common difficulties students have in learning limits as fundamental bases for designing this instruction. Basing our activity draft on what we learned from research helps us to avoid common pitfalls that have been unearthed in the past. For example, students may believe a sequence cannot reach its limit or might see the limit as a bound for the sequence (Davis & Vinner, 1986). Conceptions of limit such as these can be indicative of a larger issue that researchers have identified, that students have great difficulty in learning the formal definition of limit (Davis & Vinner, 1986; Roh, 2008; Tall & Vinner, 1981; Williams, 1991).

If students are not learning the formal definition of limit, then what informal notions are they working with? This question is one Tall and Vinner (1981) struggled as they established a distinction between what they call *concept definition* and *concept image*. Tall &Vinner call a concept definition of a concept, "a form of words used to specify that concept" (p. 2). In our case, we can consider the formal definition of limit of a sequence a concept definition. A concept image then, "describes the total cognitive structure that is associated with the concept" (p. 2). For example, if a student believes a sequence cannot reach its limit, this would be part of the student's concept image of limit of a sequence. In this distinction, Tall and Vinner outline the very means by which students can be given a rigorous, mathematical definition of a concept, and yet act with that concept in a manner counter to the definition, by building a concept image that does not align with a formal definition.

One way limits are often introduced informally is through dynamic language. For instance, we might say a sequence has a limit L if the terms of that sequence are getting "close to" or "closer and closer" to L, or if the sequence is "approaching" L. However, as researchers have found, these phrases do not carry the formality of arbitrary closeness necessary for a formal definition of the limit of a sequence (Monaghan, 1991; Oehrtman, 2008; Roh, 2008; Tall & Vinner, 1981; Williams, 1991) and therefore students' concept images will not align with a formal definition of limit. Such informal, dynamic views of limit can be productive for introductory calculus courses that deemphasizes limit as an overarching concept, but will not satisfy educators looking to encourage a deeper understanding of both limit and calculus.

This leaves us with a conundrum, if students are not being successful at learning the formal definition of the limit of a sequence and their informal notions can be potentially counter-productive for learning a formal definition of limit, how should instructors approach teaching limit of a sequence? A potential answer lies in combining these approaches: structure activities on limit of a sequence that help students develop informal notions that align with a formal definition. For example, with the development of his "approximation framework," Oehrtman (2008) showed that students "often naturally reason about limit concepts in terms of approximation in ways that are structurally equivalent to…epsilon-N definitions" (p. 72). In his experiments, Oehrtman showed how students have intuitive notions of bounding error terms that closely align with the arbitrary closeness necessary for the formal definition of limit of a sequence. Similarly, Roh (2008) designed a manipulative called "epsilon strips" that students used to judge the appropriateness of various definitions of limit. What each researcher has shown is that students can build a concept image of limit of a sequence that is structurally similar to a formal definition.

Fundamentally we agree with Oehrtman (2008) that an introductory activity for limit of a sequence should "lay conceptual groundwork from which formal understandings may emerge but not necessarily…provide those formalizations themselves" (p. 70). We take this as our primary objective in creating the "Getting Back to Sesame Street" activity: Students' informal views of limit of a sequence, which are components of their concept images, should form a foundation for a formal definition of limit of a sequence. Furthermore, in order to help students avoid the potential conceptual trappings of the informal, dynamic view of limit, we aim to design the "Getting Back to Sesame Street" activity so that it will "induce dynamic images that are compatible with the [formal] definition of limit" (Roh, 2008, p. 231). This goal of guiding students' concept images to align with a formal definition guides the design and implementation of the activity.

Phase two: Broad learning goals

Phase two of this first design cycle focused on using the research on literature to outline the broad learning goals for the activity focusing on the limit of a sequence. These goals captured aspects of the learning goals for both the current activity under development as well as the course at large. Our original five broad learning goals for the activity are shown in Figure 1; a short discussion of each is also warranted.

1. Students will see calculus presented in a way that connects it to elementary mathematics.
2. Students' views of infinite sequence will be grounded in physical experience.
3. Students will investigate the concept of a sequence via an activity based on Zeno's paradox.
4. Students will develop an informal view of limit of a sequence that "matches" the formal definition.
5. Students will be introduced to the formal definition of the limit of a sequence.

Figure 1. Phase one broad learning goals.

Goal #1 is the primary goal of the CELTIC curriculum development project. We not only want our preservice elementary teachers to learn calculus, but for them to see calculus as being connected to the elementary mathematics they will teach. In order to satisfy this goal, our activity would have to present the limit of a sequence in a manner that showcases a meaningful connection to elementary mathematics. Goal #2 was included because we felt that one way we could help students have a meaningful experience was to connect their mathematical work to a physical experience. This goal would be realized if the activity engaged the students to interpret the mathematics in light of their physical experience. Goal #3 speaks to the foundation of the activity under development. Often times, investigating the manner under which a concept was understood historically can provide interesting fodder for classroom activities. In this case, our activity engages students to think about the limit of a sequence in a manner similar to the way the Greek's handled such concepts in antiquity. Goal #4 was directly informed by the work of Oehrtman (2008) and Roh (2008). In both researchers' work, it was apparent that students are often able to be successful in describing complex limiting behavior informally. Furthermore, Oehrtman stresses that instructors should aim to align these informal conceptions with the formal definition. This activity goal stresses that importance by establishing that our activity should lay such a foundation for students so that their informal notions of limit are applied in a manner that aligns with the formal definition the students will be required to learn. Finally, Goal #5 was included because we felt students can be guided to not only describe limiting behavior informally, but that they should learn to address the formalization of such behavior. Given that in traditional calculus courses, this definition is often simply given to the students, we felt students should be able to bring their informal notions of limit in line with the formal definition, which would be presented using the language of the activity.

Phase three: Activity draft

The activity is composed of two parts, one preliminary part that helps students visualize the process under discussion and the featured part in which students investigate an infinite sequence and that sequence's limit. The entirety of the activity draft has been included in a condensed manner (see Figure 2); spaces for student work and some images have been omitted for length considerations. Students will work on this activity in groups of size four to five.

Phase four: Small group pilot of activity

Setting and Participants

In Spring 2011, the research team piloted the activity draft at a large research university where the instructional materials are being developed. Elementary education majors were contacted via email and offered a modest stipend to participate. Four volunteers chose to participate and these four were split into two pairs, herein referred to as Group 1 and Group 2. Each pair of students was asked to work through the activity with the guidance of one or two graduate research assistants acting as instructors. The students were mainly left to their own devices unless specific instruction was required or requested. The sessions were video recorded.

The focus of the analysis of these sessions, undertaken both individually by the members of the research team as well as collectively as a group, was two-fold: to analyze the activity primarily by assessing how successful students were in meeting the activity's learning goals and to analyze the learning goals themselves. In viewing the sessions, the research team discussed and made notes of how well the students were able to answer the questions, the clarity of question prompts (e.g. did students answer any latter questions within earlier ones, thereby indicating the latter would be unnecessary), and overall engagement. These discussions and notes were later used to make edits to both the learning goals and the activity.

Getting Back to Sesame Street: An Introductory Activity for the Limit of a Sequence

Big Bird and Count von Count are traveling back to Sesame Street when they come to a bridge. Just before the bridge there is a warning sign: "Lord Zeno's Bridge! You can only cross this bridge by counting your steps!" Big Bird says to Count von Count, "Count, you're a counting genius! I know how to get across Lord Zeno's bridge, you count my steps and I'll walk across!" Unfortunately, when they get closer, Big Bird notices another sign: "Each step on my bridge must be special: Every step you take must be exactly half of the remaining distance you have left to cross." Lord Zeno must be one odd fellow; this is a strange bridge indeed! Let's try and figure out how Big Bird and Count von Count will get across!

TASK #1: Try it out!

Go and find a spot about 5-6 ft. from a wall. Try to follow the directions on the sign: Every step you take should be half the distance that is between you and the wall. Keep taking steps until you can't take any more! Once you are done, sit down with your group and answer the following questions:

1. How many steps did you take?
2. Why did you have to stop taking steps?
3. Could you have taken more steps? If yes, explain what would have to change in order to do so. If not, explain why your number of steps is the maximum.

TASK #2: Big Bird & Count von Count

Let's help Big Bird and Count von Count figure out just how many steps Big Bird will have to take in order to cross this strange bridge.

1. Let's say that Zeno's bridge is 10 meters long, how many steps do you estimate Big Bird will have to take in order to finish crossing the bridge? Explain your estimation.
2. How far will Big Bird travel with his first step? His second?
3. Construct the first five terms of a sequence by computing the lengths of each of Big Bird's steps.
4. Would Big Bird ever have a step size less than .001 meters? If yes, which step would that be? If not, why not? Explain your reasoning.
5. Without computing, do you know if Big Bird will ever have a step size less than .000000001 meters? How about 10^{-100} meters? How could you find the number of steps?
6. What is the smallest distance Big Bird will ever cover with one step? Explain how you know.
7. Mathematicians use the concept of a "limit of sequence" to describe sequences' long-term behavior. In the story, you have created a sequence of Big Bird's step lengths that goes on and on. If we call the length of the n^{th} step s_n (e.g. Big Bird's 1^{st} step would be s_1, his second step would be s_2 and so on) then we can write the following:

$$\lim_{n \to \infty} s_n = ?$$

8. What do you think this limit equals? Discuss the limit notation with your group.

Figure 2. Condensed first activity draft.

Results

In Part #1, both groups took about four steps before deciding they could not take any more. When asked why they had to stop taking steps, individuals from both groups commented that they were "too close to the wall." Group 2 decided that if they had smaller feet they could have taken more steps since, "technically we could get closer and closer but never approach the wall." Group 1 decided that they could not have taken more steps, but their reasoning was similar to that of Group 2: "the wall stops us no matter what, and if we keep splitting then we'll never reach zero," and "no because we can't go through the wall/our feet cannot get any smaller." In both groups, the issue of foot size was seen as relevant to the matter of continuing to take steps as described in the activity. However, it is noteworthy that each group demonstrated some understanding of the ability to split a length indefinitely.

In the beginning of Part #2, students were asked to estimate how many steps Big Bird will need to take in order to cross this strange bridge. Group 2 claimed five steps would be enough as one of the members reasoned, “He will be so close to the end by that point that his feet will cross over.” However, the other member of Group 2 expressed “he will never actually be able to cross cause he won’t reach zero.” It is obvious that the infinite process is attended to by this group, but they end up concluding the discrepancy is handled by the fact that his feet will cross over once he gets very close. Group 1 decided Big Bird would take an infinite number of steps, stating, “Big Bird will never get to the end of the bridge, but he will not be able to take small enough steps at the end.” Throughout Part #1 and into Part #2, students reasoned using the physical act of crossing the bridge.

In questions #2-3, students were asked to compute the distances of Big Bird’s first five steps and in question #4, they had to consider whether Big Bird would have a step size less than 0.001 meters. Both groups were successful in computing these figures and both groups used a calculator and repeated halving to arrive at the correct answer of 14 steps for question #4. Interestingly, before computing, one member of Group 2 commented that “technically, yes” Big Bird would have a step size less than 0.001 meters showing support that this student could conceive of a sequence in which the terms are getting arbitrarily close to zero.

Question 5 asked, “Without computing, do you know if Big Bird will ever have a step size less than 0.000000001 meters? How about 10^{-100} meters? How could you find the number of steps?” This question posed a serious challenge for both groups. Group 1 spent a considerable amount of time trying to come up with a closed form for the sequence. One of the members had learnt this material and remembered aspects of what the formula would look like, but could not immediately remember the exact formula. It was observed that in the context of the course, the activity would be situated after a unit on sequences and that the interview participants’ background knowledge would be limited as compared to those students in the course.

In the final question concerning the limit definition, Group 1 found that the limit of the sequence was zero, but Group 2 was not able to make sense of the limit definition and each group required rather extensive discussion with the instructors. It was apparent this question was extremely difficult for the students.

Phase five: Revisions

Revisions to the Learning Goals

After analyzing the groups’ performance on the activity, the research team revisited each learning goal to assess whether it was appropriate for the activity. Only Goals #1 and #3 were changed. Goal 3 initially read: “Students will investigate the concept of a sequence via an activity based on Zeno’s paradox” and in its revised state reads: “Students will investigate the concept of a sequence *with limit zero* via an activity based on Zeno’s paradox.” The added emphasis on the limit of the sequence being zero is an attempt to be more specific in the learning goals.

The edit to Goal #1 signifies a more significant alteration to the research project as a whole. Initially, the goal read: “Students will see calculus presented in a way that connects it to elementary mathematics” and after editing reads: “Students will see that calculus can be learned in an elementary context.” This change was a result of conversations among us and with others concerning the difference between using actual elementary-level mathematics in the CELTIC activities versus simply using an interesting setting for a calculus activity. In this case, it was apparent that the addition of a Sesame Street setting was not connecting the mathematics of calculus to actual elementary-level mathematics, but more engaging students in a fun situation.

Revisions to the Activity

Overall, the interview results were positive indicators that the activity was successful in helping students learn, informally, what mathematicians mean by the limit of a sequence. The most significant difficulty we witnessed was in students' understanding of the final question where the formal definition of limit of a sequence was introduced and the students were asked to decide on the value of the limit of the sequence they had created.

Part #1 remained unchanged. The research team felt strongly that having the physical experience of continually halving a distance was useful for the students in conceptualizing the behavior and successfully completing the activity. While students in the interviews seemed to focus on the physical limitations of taking steps that were getting extremely small, each group was able to consider the more abstract behavior of infinitely halving with the concrete limitations of step size and these conversations seemed productive.

Most of Part #2 also remained unchanged until the final few questions, leading up to the formal definition of limit of a sequence. In order to make students' transitions from the informal notions of limit to the formal definition more successful, a question was added to deal with the idea that given any arbitrary small distance, Big Bird will eventually take a step smaller than the given distance. The question was worded, "Big Bird makes a shocking revelation: He claims that if you call out any number, as small as you like, if he follows Lord Zeno's directions, after a certain step, the size of all his following steps will be smaller than your number. Test out Big Bird's theory."

Finally, since students had a difficult time understanding the definition of a limit of a sequence as it was presented in Part 2, the language was simplified. The research team decided the description of the limit of a sequence was too long and confusing, thus it needed to be pared down. Figure 3 contains the revised wording of the final question including the limit definition.

> Mathematicians use the concept of a "limit of a sequence" to describe sequences' long-term behavior. The idea behind a limit of a sequence is a bit different than a speed limit or credit card limit. We say a sequence has a limit (let's call that limit L for now) when we can show that for any incredibly small number you can imagine, we can find a spot in our sequence such that for every term after it, the difference between L and any of those terms is less than the incredibly small number we picked. Do you think your sequence for Big Bird's steps has a limit based on the description given above? What do you think that limit is?

Figure 3. Question #9 from revised activity draft.

Design Cycle Two

The second design cycle for the "Getting back to Sesame Street" activity started with the piloting of the revised activity draft. The pilot took place in the experimental "Calculus for Elementary Teachers" class as part of the first implementation of the new course. Following the activity pilot, the team analyzed the data collected in order to make more refinements of the goals and activity, as will be discussed in the results section. The cycle ended with the creation of a new version of the learning goals and the current version of the activity to be used in the next implementation of the course.

Phase one: Class Pilot of Activity

Setting and Participants

In the fall of 2011, the initial realization of the new course Calculus for Elementary teachers occurred; 29 students enrolled in the course. The course met twice a week for 75 minutes each. The students were recruited from the incoming freshman cohort of elementary education majors. Students were originally placed in the standard calculus course, and they were recruited to enroll in the experimental course. The advisor for the students, as well as the department chairperson for Elementary Education conducted the recruitment. The course was team taught by two of the researchers and supported by the third researcher. Instructors in the experimental calculus course emphasized the big ideas of calculus as well as modeled the teaching strategies that they hope will be implemented by the future teachers. Both teachers were in the classroom all the time, but they traded off the facilitation of the class. The instructional strategies used include: inquiry, collaboration, active learning, justification of ideas, guided whole class discussion, and provision for diverse learners. The third researcher took detailed field notes and occasionally video recorded the class period as needed.

Data Collection

In general, we decided to videotape only the class days when we were implementing introductory and investigatory activities. On the days that the students did the activities, we videotaped one small group and the whole class discussion. The implementation of the "Getting Back to Sesame Street" activity occurred on the sixth and seventh class meeting of the semester. The students' work from all seven groups was collected. With the five learning goals for the activity in mind, the team watched the videos together, analyzed students' work, and noted student discourse that confirmed or disconfirmed how the students' participation showed their academic progress.

Results of Analysis: Positive general conclusions

Engagement with the physical situation. First of all, we observed that the students were physically and intellectually engaged in the activity. The observations and the student work that was collected indicate that the students found the activity interesting and worthwhile. They were interested in the issue of walking to the wall that Part 1 involved. The question, "Why did you have to stop taking steps?" elicited several different answers; the most common answer was a variation on the idea that smaller feet would help generate more steps.

Prior learning. The second positive result of the activity regards prior student learning. During the original interviews from Design Cycle 1, the students did not have earlier instruction on sequences and writing recursive and closed forms for sequences. This was not the case during the class implementation, most of the students did not have any difficulty working with the abstract notation and developing the analytic forms as asked in Part 1, Question #6. This shows that the earlier instructional sequence worked well in support for this activity.

Relating mathematics to a real situation. Finally, we found that students were able to struggle with, connect, and resolve some of the issues of relating a "real" physical situation to the mathematics found in the instructional activity. One common remark concerned the "technical" issue and the "practical" issue, where students offered their thinking about what would happen in each case. Several students would say, "Well, technically they would never get there…" indicating that they made a distinction between what would happen in the context described by the activity and what would actually happen. This idea pervaded the second part of the activity as well. We think students' informal dealings with this difference can be used to build deeper understanding of the theoretical idea of limit. Furthermore, students did find the issue of getting smaller and smaller somewhat paradoxical, and talked about why Big Bird could not just "step off the bridge".

Results of Analysis: Conclusions about concerns

Computation. Not surprisingly, the students in all but one of the groups struggled with finding the exact N that would yield the closeness that was asked for (how to get within k of the limit). This computation had been deliberately constructed so that trial and error would not be appropriate, so using logarithms was necessary to solve. The great majority of the students had learned how to solve an equation by taking the "log" of both sides, but could not remember how to do it. Because they could not do the computation, they shifted their focus from the big picture to how to "do the problem." The instructors ended up spending considerable instructional effort on re-teaching logarithms and algebraic manipulations, weakening the effect of the activity on supporting the idea of "arbitrary closeness" that the question was meant to address.

Limit zero obscuring limit definition. After analyzing the video and student work, we determined that the idea of looking at the difference between a term of the sequence and the limit becoming arbitrarily close was obscured for many of the students. Even when looking at the step size getting smaller and smaller, students did not seem to indicate 0 as a limit. More significantly, even if they did, that did not promote thinking about the notion of getting "as close as we want" to some value, one of the phrases that the researchers believe is structurally similar to the formal definition.

Understanding infinite processes. We considered the concept of infinity and how the students used their understanding of infinite processes, and found that our students seemed to have little experience to call on when needing to reason about infinite processes. For example, the notion that the terms of the sequence generate forever but can still "end" somewhere seem to be contradictory for the student.

Premature introduction of formal definition. One observation that we made for all of the small groups is that the students could not understand and articulate even an answer for the last question in the activity. The video of one of the groups showed one of the researchers actually telling the students her understanding of the last question. However, even after she did this, the students only wrote down the final answer, zero. This conclusion to the activity did not allow us to finish with support for the students to build a strong concept image, especially their informal notion of limit. This aspect of the activity was removed, and a plan to develop a follow-up activity that addresses the formal definition has been initiated.

Phase two: Revisions

Revisions to the Learning Goals

Following is the most recent version of the broad learning goals. They show a change in Goal 3 in that the limit of the sequence is changed from zero to ten and Goal 5 is removed in response to the concerns discussed above.

1. Students will see calculus presented in a way that connects it to elementary mathematics.
2. Students' views of infinite sequence will be grounded in physical experience.
3. Students will investigate the concept of limit of a sequence with limit ten via an activity based on Zeno's paradox.
4. Students will develop an informal view of limit of a sequence that "matches" the formal definition.

Figure 4. Current activity learning goals.

Additionally, in order to give students more time to familiarize themselves with the idea of infinite process and formalization of the limit of a sequence, the decision was made to develop and introduce a pre-activity on infinite processes and a post-activity connect the results of the activity with the formal definition of limit of as sequence. This will create an instructional sequence of three activities.

We conclude our description of this design cycle with a condensed version of the most recent version of "Getting Back to Sesame Street" in Figure 5 below.

Getting Back to Sesame Street: An Introductory Activity for Limit of a Sequence

Big Bird and Count von Count are traveling back to Sesame Street when they come to a bridge. Just before the bridge there is a warning sign: "Lord Zeno's Bridge! You can only cross this bridge by counting your steps!" Big Bird says to Count von Count, "Count, you're a counting genius! I know how to get across Lord Zeno's bridge, you count my steps and I'll walk across!" Unfortunately, when they get closer, Big Bird notices another sign: "Each step on my bridge must be special: Every step you take must be exactly half of the remaining distance you have left to cross." Lord Zeno must be one odd fellow; this is a strange bridge indeed! Let's try and figure out how Big Bird and Count von Count will get across!

TASK #1: Try it out!

Go and find a spot about 5-6 ft. from a wall. Try to follow the directions on the sign: Every step you take should be half the distance that is between you and the wall. Keep taking steps until you can't take any more! Once you are done, sit down with your group and answer the following questions:

1. How many steps did you take?
2. Why did you have to stop taking steps?
3. Could you have taken more steps? If yes, how many more steps could you have taken? If not, explain why your number of steps is the maximum.

TASK #2: Big Bird & Count von Count

Let's help Big Bird and Count von Count figure out just how many steps Big Bird will have to take in order to cross this strange bridge.

1. Let's say that Zeno's bridge is 10 meters long, how many steps do you estimate Big Bird will have to take in order to cross the bridge? Explain your estimation.
2. How far did Big Bird travel with his first step? How far has he traveled after his second step?
3. Let $\{a_n\}$ be the sequence where the n^{th} term corresponds to the distance Big Bird has traveled after his n^{th} step. Compute the first five terms of this sequence.
   (Hint: You computed a_1 and a_2 in question #2).
4. Would Big Bird ever be within 0.5 meters from the end of the bridge? If yes, give support for your answer. If not, explain why not.
5. Without computing, do you know if Big Bird will ever be within 0.001 meters from the end of the bridge? How about 0.00001 meters? How do you know?
6. Assuming that Big Bird gets within 0.001 meters from the end of the bridge, how far might he have traveled in total? What about if he gets within 0.0001 meters of the end of the bridge?
7. Big Bird makes a shocking revelation: He claims that if Count calls out any distance, as close to 10 as he wants, if he follows Lord Zeno's directions, after a certain step, he will have traveled that far. Do you believe this is true? Why or why not?
8. You may have noticed that the distance Big Bird has traveled seems to be getting close to 10. In your group, write a description of the behavior that the sequence exhibits as Big Bird continues his journey to the end of the bridge.

Figure 5. Condensed activity – current version

Conclusions

The work of this research project is an example of design research where experts design instruction (Brown, 1992). Research that has been conducted and reported in the last twenty years is an important base to build on, and we certainly found that the research of our colleagues valuable and also applicable for the students in the Calculus for Elementary Teachers course. However, there is no substitution for the actual implementation and testing of the instructional materials in a real setting with students; thus, the interviews and piloting of the materials played a pivotal role in our material development process. In this paper, we only reported on the research for one of the instructional activities, the introduction of limit of a sequence. This kind of research is particularly time intensive as we are attempting to do this kind of analysis and design cycling for all the activities we have developed. In design research for a whole course development, this work would happen for the guided lecture and homework as well, but time will not permit that in this case.

Finally, the importance of the work of training STEM-focused elementary teachers cannot be emphasized enough. In our society, elementary teachers have often been seen as not liking or wanting to teach mathematics and science. This needs to become a priority: for our future teachers to infuse value and excitement into the math and science learning of children at the elementary level. When this happens, we can expect improvement in mathematics and science student learning in middle school, high school, university, and in the workplace. The philosophy of Calculus for Elementary Teachers is that bringing preservice teachers into the mathematics communities as engaged learners and connecting their learning to their future work as teachers may become one of the primary venues for changing the numbers and quality of our future STEM professionals.

References

Ball, D.L., Hill, H., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, Fall 2005*, 14 - 46.

Ball, D.L., Thames, M.H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*, 383-388.

Brown, A. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences.* 2(2). 141-178.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education* 23, 2-33.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schaubel, L. (2006). Design experiments in educational research. *Educational Researcher, 32*(1), 99-13.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues, *The Journal Of The Learning Sciences*, 13(1), 15-42.

Common Core State Standards Initiative. (2010). Common Core State Standards for Mathematics. Retrieved from http://www.corestandards.org/

Davis, R. B. and Vinner, S. (1986). The notion of limit: Some seemingly unavoidable misconceptional stages. *Journal of Mathematical Behavior, 5*(3), 281–303.

Fennema & Franke (1992), Teachers knowledge and its impact. In Grouws, D. Ed. *Handbook of Research on Mathematics Teaching and Learning*. Reston: NCTM.

Hill, H., Rowan, B., & Ball, D. B. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Mamona-Downs, J. (2002). Letting the intuitive bear on the formal: A didactical approach for the understanding of the limit of a sequence. *Educational Studies in Mathematics, 48*(2-3), 259-288.

Mewborn, D.S. (2001). Teaching, teachers' knowledge, and their professional development. In J. Kilpatrick, W.G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 45-52). Reston, VA: NCTM.

Monaghan, J. (1991). Problems with the language of limits. For the Learning of Mathematics, 11(3), 20-24.

National Mathematics Advisory Panel (2008). *The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

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 Teaching in Undergraduate Mathematics Education*. Washington, D.C.: American Mathematical Society.

Papert, S. (1971). "Teaching children to be mathematicians vs. teaching about mathematics." (Al Memo No. 249 and Logo Memo No. 4). Cambridge MA: MIT.

Phillip, R. A. (2007). Mathematics teachers' beliefs and affects. In F. K. Lester (Ed.) *Second handbook of research on mathematics teaching and learning.*

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.
Roh, K. (2008). Students' images and their understanding of definitions of the limit of a sequence. *Educational Studies in Mathematics, 69*(3), 217-233.
Seaman, C. E. & Szydlik, J. E. (2007). Mathematical sophistication among preservice elementary teachers. *Journal of Mathematics Teacher Education, 10*, 167-182.
Stylianides, G. & Stylianides, A. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*, 314-352.
Tall and Vinner (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151-169.
Thompson A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.) *Handbook of research on mathematics teaching and learning*. Reston: NCTM.
Tudge, J.R. H. & Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In M. H. Bornstein & J.S. Bruner (Eds*.), Interaction in human development* (pp. 17-40); Hillsdale NJ: Erlbaum.
Williams, S. (1991). Models of limit held by college calculus students. *Journal for Research in Mathematics Education,* 22(3), 219-236.
Wu, H. (2006). How mathematicians can contribute to K–12 mathematics education. In *Proceedings of the International Congress of Mathematicians,* Madrid 2006, Volume III, European Mathematical Society, Zurich, 1676-1688.

STUDENT UNDERSTANDING OF SYMBOLS IN INTRODUCTORY STATISTICS COURSES

Hyung Kim
The University of New Hampshire
hwy3@wildcats.unh.edu

Tim Fukawa-Connelly
The University of New Hampshire
Tim.FC@unh.edu

Samuel A. Cook
Wheelock College
scook@wheelock.edu

This study explores student understanding of the symbolic representation system in statistics. Furthermore it attempts to describe the relation between student understanding of the symbolic system and statistical concepts that students develop as the result of an introductory undergraduate statistics course. The theory, drawn from the notion of semantic function that links representations and concepts seeks to expand the range of representations considered in exploring students' statistical proficiencies. Results suggest that students experience considerable difficulty in making correct associations between symbols and concepts; that they describe the relationship as seemingly arbitrary and that they are unlikely to understand statistics as quantities that can vary. Finally, this study describes students' need for robust knowledge of preliminary concepts in order to understand the construct of a sampling distribution.

Keywords: statistical symbols, symbolic representation, symbolic fluency, introductory statistical concepts

Research Questions

In the field of mathematics, significant importance was placed upon symbolic representations of communication, teaching and learning (Arcavi, 1994). In particular, students at introductory level statistics courses have been found to mix up the symbols for statistics and parameters (Mayen, Diaz & Batanero, 2009), which could hinder them from developing the concepts that such symbols represent. However, our literature search suggests that there have not been any studies published that explore students' understanding of the symbolic system of statistics. Therefore, we investigate the following questions:

- How do students perceive the symbols for mean and standard deviation after a lecture course?
- How does students' symbolic fluency relate to their ability to make sense of more advanced statistical concepts?
- When students have a strong mathematical background, how does that support or inhibit their ability to be successful in developing symbolic reasoning in statistics?

Previous research suggests (Mayen, Diaz & Batanero, 2009) and our results confirm that students find the choices of symbols arbitrary and difficult to associate with related concepts, and that students need particularly strong conceptual and symbolic understandings in order to make conceptual sense of the standard deviation of a sampling distribution. We also found that student understanding of the relation of *statistics* to *parameters* was not robust, and they did not consistently view statistics as variables. We found that many students did consistently look for meaning based upon the symbolic representation of concepts.

Literature Review

Onto-semiotic research proposes that "representations cannot be understood on their own. An equation or specific formula, a particular graph in a Cartesian system, only acquires meaning as part of a larger system with established meanings and conventions" (Font, Godino, & D'Amore, 2007, p. 6). The implication is that the system of practices is complex in that each one of the different object/representation pairs provides, without segregating the pairs, a subset of the set of practices that are considered to be the meaning of the object (Font, Godino, & D'Amore). Within the realm of statistics, even when the object under consideration seems relatively simple, such as the mean, there are often multiple symbolic representations used interchangeably. For example, $\bar{x} = \sum x_i \big/ n$ may be used without consideration of any other type of representation: graphical, verbal, etc. The relationships between object and representations become even more complex when moving toward a more complex idea, such as the standard deviation of a sample mean. Due to a layering of representations, it is conceivable that the different possible pairs of object/representation convey different meanings of the same object.

The onto-semiotic approach requires a discussion of the role of communities of practice in order to describe the representation system, since representations (symbols) are only ascribed meaning by those who work with them. Eco (1976) gave the term ***semiotic function*** to describe the dependence between a text and its components and between the components themselves. The semiotic function relates ***the antecedent*** (that which is being signified) and ***the consequent sign*** (that which symbolizes the antecedent) (Noth, 1995). The members of the statistical community and their representation system define a complex web of semiotic functions. It is important to note that these functions "… the role of representation is not totally undertaken by language (oral, written, gestures…)" (Font, Godino, & D'Amore, 2007, p. 4).

For example, when learning the standard error of a sample mean, students are confronted with the simple looking formula: $\sigma_{\bar{x}} = \sigma/\sqrt{n}$. This formula has a seemingly simple explanation: "*the population standard deviation* of *the sample means* is given by the population standard deviation divided by the positive square root of the sample size." In this case, the representation $\sigma_{\bar{x}}$ draws on the agreed-upon symbols for *the population standard deviation* and *the sample mean* to communicate the meaning "the population standard deviation of the sample means." However, it does not give information about how to determine the value. Moreover, the symbol $\sigma_{\bar{x}}$ requires students to be able to make sense of a mixture of previously separate representational systems: those that represent statistics derived from a sample (for example, $\bar{x}$) and those that represent parameters derived from a population (for example, σ). When given the representation on the right-hand side of the equation, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, students read a formula that implies they should perform a calculation by mixing the pieces of symbols up from separate representational systems. Most importantly, the right and left-hand symbols could be interpreted as different meanings of the standard error of the sample mean. So, students are potentially confronted with various possible representations of the same object as described above.

A representation is "something that can be put in place of something different to itself and on the other hand, it has an instrumental value: it permits specific practices to be carried out that, with another type of representation, would not be possible" (Font, Godino, & D'Amore, 2007, p.7). In this case, the standard error of a sampling distribution, the object $\sigma_{\bar{x}}$, can be understood as a necessary concept (Hewitt, 1999) that emerged from a system of practices. It should be considered unique, with a holistic meaning that is agreed upon by the community of practice; however, the concept is expressed by a number of different semiotic functions. Each of these object/representation pairs should be understood as encapsulating a different possible set of meanings and enabling different practices.

Rationale

A literature search suggests that although there have been investigations of students' understanding of measures of center (Mayen, Diaz, & Batanero, 2009; Watier, Lamontagne, & Chartier, 2011), variation (Peters, 2011; Watson, 2009; Zieffler & Garfield, 2009), and even students' preconceptions of the terms related to statistics (Kaplan, Fisher, & Rogness, 2009), no one has yet explored student understanding of the symbolic system of statistics. One paper did draw upon the onto-semiotic tradition to describe student errors related to representations of the mean and median (Mayen, Diaz & Batanero, 2009).

Hewitt (1999, 2001a, 2001b) distinguished those aspects of a concept that can only be learned by being told and then memorizing, which he labeled *arbitrary*, from those that can be learned or understood through exploration and practice, which he labeled *necessary*. This distinction between symbols of the mathematical system suggests the importance of symbolic representations in building conceptual understanding and procedural fluency. Hewitt noted that names, symbols and other aspects of a representation system were culturally agreed-on convention. Although symbols may seem sensible once an individual has an understanding of the culture, "names and labels can feel arbitrary for students…there does not appear to be any reason why something has to be called that particular name"—after all, as he argues, "there is no reason why something has to be given a particular name" (1999, p. 3). Hewitt pointed out that for students to communicate with experts, they must memorize the arbitrary elements and correctly associate them with appropriate understandings of the necessary elements.

Recently, Shaughnessy called for research into "students' conceptions of the interrelationships of the aspects of a distribution" (2007, p. 999). But he focused only on the special place of graphs as a tool in statistical thinking, and did not acknowledge the importance of the representational system in which graphs are situated. The research on students' conceptual understanding of statistical concepts has avoided discussion of the importance of representation; yet, onto-semiotic research claims that descriptions of conceptual understanding are incomplete when pursued only via one or two possible representations of a concept. This study contributes to the growing body of research on student understanding of statistical concepts by describing students' symbolic fluency and the ways they link concepts and symbols.

The History of Changes in Statistics Courses

In the past 25 years, there has been significant change in the structure of the introductory statistics curriculum; where there once was a focus on learning probability, theory, and formulae, there now is a data-driven approach to content via descriptive statistics, basic probability, and inferential statistics (Garfield & Ben-Zvi, 2008). The focus of reform, especially because of recent technological advances, has been to emphasize statistical thinking. This includes using data, understanding the importance of data production, and appreciating the presence of variability (Garfield & Ben-Zvi). The statistics education community has also adopted the view that the course should "rely much less on lecturing, much more on alternatives such as projects, lab exercises, and group problem solving and discussion activities" (Garfield & Ben-Zvi, p. 12). One way to define the difference between traditional, lecture-based statistics classes and inquiry or reform-oriented classrooms is to compare how much of the responsibility for mastering cognitive processes is given to the students.

While the curricular organization of the courses in this study conformed to those typically found in a reform-oriented classroom, the instruction itself was essentially traditional. The instructors had almost total responsibility for daily classroom activities and the content was delivered primarily via lecture.

Methods

Data for this study was drawn from eight participants in a mid-sized public university in New England. Two of the participants were in a lower level introductory statistics class and six were from an upper level class. The lower level class was designed to allow first-year students to meet the general education requirement of the university, and thus is non-calculus based. The upper level class was designed to serve mathematics majors, and thus is calculus-based. The two courses occurred in the same semester.

We used a phenomenological approach to collect data, the process of which was conducted in two steps: a survey assessment and a follow-up interview. For the survey, we developed a fourteen-item assessment, which is attached in the appendix at the end of this paper. Some of these items were modified from *Assessment Resource Tools for Improving Statistical Thinking,* developed by the faculty members of the University of Minnesota in 2006. The rest of the items were created by our research team. The assessment items sought to evaluate student understanding of what the symbols represented and their conceptual understanding primarily via their symbolic representations.

The goal of the interview process was to identify how students' understanding of symbolic representations and their level of symbolic fluency potentially impacted their understanding of certain symbol-oriented concepts. The interview of the two participants from the lower level class was conducted a few days after the survey; the interview of the seven participants from the upper level class was conducted immediately after the survey. Based upon their work on the content survey the nine students appear to range from low achieving to high achieving in statistics.

Both the survey and the interview were analyzed qualitatively. All interviews were audio-recorded and transcribed. For coding, each utterance was assessed to examine the information it gave about symbolic understandings. Then, within each transcript, we categorized and summarized the utterances that deemed informative understandings by the type of concepts and connections it described with their symbolic understanding. We read within and across categories to develop conclusions. We continually rechecked our conclusions against the data that described the students' proficiencies. In this process, to find out how students' understanding of concepts in descriptive statistics is related with their ability to make symbol sense, parts of the grounded theory approach were blended in.

Results

Through the data analysis process, we drew three conclusions regarding the pedagogical difficulties that many participants encounter when attempting to reason symbolically in statistics. We also detected that high achieving students face a pedagogical hindrance caused by their academic disposition. A detailed description of the findings is illustrated below.

1. *Students find the choice of symbols seemingly arbitrary and difficult to associate with related concepts*. According to onto-semiotic research, holding various connections that a concept has with its various expressions is essential for one to internalize the concept. One of the connections is associated with the symbol that typically represents the concept. In introductory statistics courses, many concepts of descriptive statistics are introduced with their associated symbols. The choice of the symbols, however, is somewhat arbitrary and students have difficulty connecting the symbols with the concept that they represent. For example, consider the following claims made by Aaron:

Aaron: Well, it (μ) is sort of the mean of the whole population. So, it's the big mean, as opposed to the sort of small, local mean (for $\bar{x}$).

Interviewer: Okay. And then, the notation for the smaller one is …

Aaron: It seems arbitrary to me. It just seems like they didn't have a good symbol, so they just used *x-bar*. …… But it's one of those where I just remember it, because I just had to force myself to memorize that. There's no intuitive connection there, to mean. It's just, someone said that that's what that is. So that's what I remembered it to be.

…

Interviewer: Okay. What about *sigma*, there? What's your understanding of *sigma*?

Aaron: *Sigma* would be the standard deviation. The *sigmas* actually make more sense. *Sigma*, being the standard deviation, at least there's the relationship, there's *s*. So, you know, I guess, it's interesting that they used the Greek *s* for the sort of whole standard deviation, where sort of local, standard deviations have regular, lower case *s*. But in the case, like, it's more intuitive than *x-bar* for the observation.

In this example, while Aaron acknowledges the importance of the symbolic connections, he struggles to find such connections. If students do not connect the concepts with their associated symbols in descriptive statistics, they will be hindered from acquiring new concepts about inferential statistics. The items 1, 2, 5 and 6 on the survey were designed to assess students' ability to discern the symbols for statistics from the symbols for parameters. While students' responses on the assessment instrument regarding symbols were 72% correct overall, they consistently reported, during the interviews, that they struggled to understand the difference between statistics and parameters and to distinguish between the symbols. Consider further, Michael's claims:

I know μ, I just always associate μ with the mean. I wasn't really sure, I don't remember if it was in the population, if it was the mean of the population or the sample, so I just kind of guessed on that one. And, for $\bar{x}$, I think I've learned that is also the mean…

He continued, "So, μ would be, like, all the data, and then, sorted, from smallest to largest, and then divided by how many were in the sample… And then, $\bar{x}$ is, I think $\bar{x}$ is the same, it's just not sorted by smallest to largest. I'm not really sure." Based on his performance on other items, it appears that Michael knows how to calculate the mean and understands what it implies mathematically. But these are only part of a complete understanding the concept of *mean*. Another aspect of understanding the mean is the ability to pair it with the distinction between sample and population, which Michael was not able to do. Instead, he attributed an incorrect difference of meanings to the two symbols for mean. While he may be able to correctly answer questions that require calculating the mean, the lack of connection may prevent him from acquiring symbolic fluency.

2. *Students need particularly strong conceptual and symbolic understandings in order to make sense of the standard deviation of a sampling distribution*. The concept of the standard deviation of a sampling distribution was determined to be one of the most difficult concepts for students in our survey. When Ian was asked to describe what a particular symbol represents, such as $\sigma/\sqrt{n}$, Ian said, "This is the population standard deviation." He continued, "(s is) the standard deviation of our sample. I think we used s in class. I'm not sure. But we used another thing to separate, just like this, our mean in our sample. And so I thought that was what it was." That is, he understood $\sigma/\sqrt{n}$ as the sample standard deviation even though the class had used s as the symbol for the sample standard deviation. This implies that he was so unsure in his knowledge that he was willing to believe that a different symbol could be substituted for s and still mean the same thing. Moreover, Ian's responses to the questions were initially definitive; only after further questioning did he admit having any insecurity of his knowledge. Even then, he did not express concern about mixed understandings or possible misattribution of meaning to symbols. We have two more examples that show students' disconnected understanding on the concepts regarding standard deviation. One of them can be seen in the case of Riley as follows:

Interviewer: But what kind of thing can we pull out, from σ and s? Does s estimate σ? Or does it estimate any of these things in here?

Riley: s over square root of n estimates σ, I believe.

Also one of our participants, Andrea, was doing very well in her class and had a very firm understanding of statistics and parameters as was shown in the following conversation:

Interviewer: Could you explain what your understanding … (is about parameters and statistics?)

Andrea: A parameter is just a piece of information about an entire population, and a statistic is a piece of information about the sample, and maybe a statistic is kind of, you use it to kind of guess at the parameter.

Further, when discussing item one, her misconception between sample standard deviation and the standard deviation of a sampling distribution was detected:

Andrea: But I kind of thought these, I had trouble, on my last exam, with, like, the difference between this one and this one. Because, like, I had a problem with --

Interviewer: The sigma over radical N, and S.

Andrea: Yeah. Because I kind of thought, I don't really, I guess I don't know what the difference, because I thought we, in class, we kind of used this to talk about the variability in a sample, but I thought S described the variability in a sample. So, I think I've got those two things kind of confused.

She acknowledges herself that she is confused with the difference between the symbols $\sigma/\sqrt{n}$ and s. We confirmed this again in the following part of the interview on item 14:

Interviewer: σ? And what is σ over radical n, then? What's the place for that? Why do we ever consider σ over radical n?

Andrea: Well, maybe I, what I thought, maybe, was that, sometimes you know what the, **maybe you know what σ is, but you don't know what that ($\sigma/\sqrt{n}$) is, and you use σ** over radical n **--**

Interviewer: You mean, we know, we don't know σ?

Andrea: Maybe, if you do know, I don't know in what situation you would know sigma but you wouldn't know s. But maybe you can use this to estimate that one?

Interviewer: You can use sigma over radical n to estimate s?

Andrea: I don't really know what I'm talking about. [LAUGHTER] But that's my best guess.

After she understood the meaning of ".. is an estimator of ..", she made a comment (in bold above) to imply that σ over radical n estimates s. One way to explain this misunderstanding is to realize that students are trained to distinguish statistics from parameters through in-class learning. Once students establish the distinction, they habitually try to discern statistics from parameters; yet their work shows that they admit to struggling in doing this. It should be noted that the expression $\sigma/\sqrt{n}$ has a great potential to confuse new learners because the symbol σ represents a *population* standard deviation, but the process of dividing by radical *n* is associated with a *sample*. Students can be easily confused as to what $\sigma/\sqrt{n}$ is associated with because they are trained to distinguish samples from population in order to be able to distinguish statistics from parameters.

3. *Students had difficulty viewing statistics as a variable.* One of the items was designed to find out if students were able to view statistics as variables and parameters as fixed constants. This skill is an essential aspect of understanding the relationship between statistics and parameters and lays the groundwork for understanding the sampling distribution. We found that all eight students had difficulty holding this view. For example, Michael said,

> I think a statistic is a calculated value, and a parameter is a, like a, it would be like a boundary that satisfies a value. *S*, so, I think s would be a, I think s would be a parameter, because sigma is the statistic. Its [measured estimator?]

Also, Brian said, "because it (s) is representative of standard deviation. I guess that varies, but—." When he was asked for the question from interviewer, "Have you thought of $\bar{x}$ as a variable before?", he answer was "No. I thought it's more just a sample, as a value that you give to a particular group." Another example is from Ian. He said, "I didn't understand that at all. I didn't know what we were looking at as, what was changing and what wasn't changing." However, with some guidance during the interview, some students were able to understand how a statistic could be viewed as a variable. For example, Andrea said, "Okay. Well, I guess, I really don't know, but I guess, my guess would be that, maybe, it would be $\bar{x}$ and s, because maybe mu and sigma don't vary, because they, I don't feel like I'm interpreting this question correctly, but I think that would be my guess, because maybe mu and sigma don't vary." These examples imply that without interruption students' understanding of statistics as a variable was minimal or nonexistent.

4. Mathematically strong students experienced special kinds of struggles in learning statistics. One of the research questions was to identify how students with a strong mathematical background develop symbolic reasoning in statistics. Thus we designed three items (4, 8, and 11) in the survey to evaluate students' reasoning level of mathematical concepts. Some participants showed strength on the algebraic and probabilistic reasoning that underlies statistical formulas. This strength was first detected via the survey and was confirmed during the interview process. For example, with the three items in the survey, while the average achievement rate of all eight participants for those three items was 61%, Ian had 100% and Jen 89%. Especially, Ian proved to have a firm understanding of the concepts focused on in the three items during the interview. For example, item 8.a in the survey asked:

> *In a university, 75% of the students are male and 25% are female. 5% of the male students and 15 % or female own a car. For each statement, determine whether it is true?*

a. We can conclude that 20% of the students in the university own a car.

During the interview, he claimed, without doing the real calculation, “I would say it's between 5 and 15. Probably around 7%?” Not only was he one of the few students who could correctly describe both the process and concept of a weighted average, but also he was able to give an approximation of the average using the four numbers shown in the question. Ian further proved his mathematical strength with his academic record showing high grades in multiple advanced undergraduate mathematics courses.

One of the characteristics that students with this disposition had was symbolic fluency. Ian, in discussing item 11, claimed:

> The center would still be zero. But the standard deviation would be sigma, because you forgot to divide. …… Because if you divided, if you do the shift first, by μ, you're centering it at zero. But if you divide x by μ first, then subtract μ, your center would actually be [UNINTELLIGIBLE], because you're going to decrease your center when you divide by μ, and then you're going to shift it the original shift. (*)

This remark of Ian’s about *z-score* shows that he understands the mathematical concepts that underlie the *z-score* formula. In this remark, it is also evident that Ian has a strong mathematical symbol sense. He was able to describe each of the pieces of the formula in terms of its relationship to function transformation; he described shifts (translations) as happening when subtracting a constant and noted that not dividing by n has no effect on the location of the center. This development of symbolic fluency (or symbol sense?), we suppose, might be the result of Ian’s pedagogical disposition because such a disposition help students to make sense of the underlying concepts of a statistical expression that use various symbols. Thus this disposition of a student would work as a great pedagogical tool for the student when explanations of statistical expressions are provided to his or her satisfaction.

However, when these mathematically strong students attempt to bring the tools that helped them be successful in K-16 mathematics to their statistics classes, they could feel as though there were different norms for perceiving mathematical concepts in statistics classes because in these classes, contrary to other mathematics classes, it is not common for instructors to provide a complete description of the statistical expressions. As such, participants claimed during the interview that the mathematical concepts were not fully explained in their classes. For example, Ian said, “I feel like we just didn’t get any of the foundational stuff. Like, this is the most lost I've ever been in a class.” He further claimed during the discussion of item 7:

> And then, there was another question where, you said, like, which of these can be considered variables, or something? Well, I never understood, he never specifically said that, and I never grasped what variables were considered, in stats. So, I guess, when you don't have that basic, basic stuff, it's, everything that comes after, you just struggle to try to put pieces together, all at the same time.

This remark not only shows Ian’s frustration that they didn’t learn basic statistical concepts from which they can develop more advanced concepts, but also reflects the conflict with Ian’s pedagogical disposition to seek out an explanation. Now, it seems as though this pedagogical disposition of Ian’s may have hindered him from developing symbol senses needed to perform well in their class reflects. For example, Ian said:

> So, now, I'm questioning myself. This median, capital *M*, is that the median of the whole population? Like, can they have the median of the sample? I've never heard that.

On one hand, such a deep understanding of statistical expressions and symbolic fluency described above in (*) was the result of the kind of academic disposition that Ian had. But, on the other hand, this academic disposition causes pedagogical conflicts with these students because they feel that the explanations provided are not to their satisfaction.

Discussion

Students, in introductory statistics courses, often struggle with symbols and making sense of concepts in relation with symbols. In an attempt to elucidate the issue, this paper addressed the following research questions:

- How do students perceive the symbols for mean and standard deviation after a lecture course?
- How does students' symbolic fluency relate to their ability to make sense of more advanced statistical concepts?
- When students have a strong mathematical background, how does that support or inhibit their ability to be successful in developing symbolic reasoning in statistics?

In investigating the first of the three research questions above, we found that the majority of students made good sense of the basic statistical symbols in descriptive statistics and distinguished the symbols for statistics from those for parameters. However, some students found the choice of symbols seemingly arbitrary and some students had difficulty associating with related concepts and attributed that difficulty to the arbitrary choice of symbols. To alleviate these difficulties, it might be necessary, as a future study, to investigate if it might be necessary that statisticians develop more systematic symbols for novices.

The second research question inquired how students' symbolic fluency relates to their ability to make sense of more advanced statistical concepts. Even though the majority of students were successful in pairing up the symbols for the mean and the standard deviation to the meanings they represent, students, in general, had trouble making sense of more advanced statistical concepts that use those symbols. In particular, it was conspicuous that students did not develop strong conceptual and symbolic understandings in order to make sense of the standard deviation of a sampling distribution. Also, the failure to view statistics as a variable was clearly shown in all eight students. The problem may have less to do with the conceptual challenge of holding that view, but more to do with some students' claim that they never had a chance to think of a statistic as a variable. To help solve this issue, we suggest that instructors give more attention to the concept of the nature of statistics in relation to their corresponding parameters. It remains, as a future study, to find what kind of examples are effective in teaching and learning how statistics vary sample by sample and thus can be treated as a variable in the given context.

The last research question focused on how the academic disposition of mathematically strong students supports or inhibits their ability to be successful in developing symbolic reasoning in statistics. This was shown in Ian's case. He had an academic disposition to seek an explanation of mathematical concepts and showed, during the interview, a strong reasoning ability about the mathematical expressions that use symbols. Our speculation on this matter is that while the academic disposition that mathematically strong students have supports their study in usual mathematics courses, this disposition could cause such students pedagogical conflicts in statistics courses. This phenomenon is attributed to the fact that in traditional statistics lectures, instructors do not provide a complete description of the statistical expressions. In order to mitigate the conflict, it would be necessary for statistics instructors to acknowledge the issue and inform students of the difference between the nature of statistics courses and that of other mathematics courses.

The findings of our paper now leave us with the following future research questions. First, at the end of an introductory statistics course, students are expected to be able to associate statistical symbols with their accepted statistical meanings and acquire the symbolic fluency. This would lay the foundation for developing a firm understanding of more advanced concepts in descriptive statistics and in the broader domain of inferential statistics. Our study suggested that, without improved practices or more instructional focus, students are likely to continue to create incorrect semiotic links and experience great difficulty in developing conceptual understanding. This leads to the next question, "what pedagogical approaches help students make better sense of symbol sense?" For example, it would be worth exploring various types of examples with which students can make better sense of symbols.

Second, we found in this study that not providing students with complete explanations of statistical concepts could hinder learning, especially for the students with the academic disposition described above. Thus the following question should be answered: "to what degree should instructors provide the explanations of statistical expressions?" Due to the dual nature of statistical concepts between mathematics and social science, it would be unrealistic to provide complete proofs of statistical expressions in class. Thus it is important to identify effective pedagogical methods that balance well between the two aspects of the discipline.

Appendix
Student Survey – Spring 2011

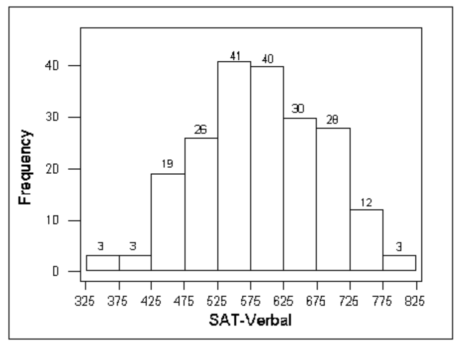
Name:

1. Select the appropriate symbol for each case:

The mean of all observations in a population. _______.
The mean of all observations in a sample. _______.
The population standard deviation. _______.
The sample standard deviation. _______.
The standard deviation of the sampling distribution. _______.

□ M s $\sigma/\sqrt{n}$ μ $\bar{x}$ σ

2. The following histogram shows the Verbal SAT scores for 205 students randomly selected from 3000 students entering a local college in the fall of 2002. Assume we know the exact values of the 205 scores.
Circle all symbols whose exact values can be found through computation using this data.



□ M $\sigma/\sqrt{n}$ μ $\bar{x}$ s σ

3. The following counts of raisins were obtained in a survey of 14 persons. These persons were selected randomly from a population of size 500.

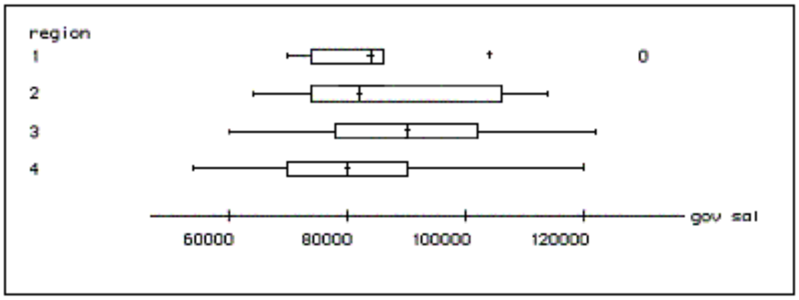
30 40 37 35 42 28 29 24 25 26 23 19 18 65

Based on the above data, please circle any of the following where an exact value can be found?

a. The mean of all observations in the population.
b. The mean of all observations in the sample.
c. The population standard deviation.
d. The sample standard deviation.
e. The median of all observations in the population
f. The median of all observations in the sample

4. Suppose that, in the above data set, the last two data values 18 and 65 are entered incorrectly. If 18 is corrected to 20 and 65 is corrected to 67, then, among the ones you circled above, which ones will change?

Items 5 to 6 refer to the following situation:
The following boxplots display the distributions of the 1993 governor's salaries according to the state's geographic region of the country. Region 1 is the Northeast, 2 the Midwest, 3 the South, and 4 the West.



5. Which region has the state with the highest governor's salary?

6. Which region has the state with the highest median governor's salary?

7. A variable is defined as a characteristic of an individual.
A variable can take different values for different individuals.

You are interested in the number of siblings that entering students of this university have. You are to take a random sample of 10 freshmen to give descriptive statistics. Let x be the variable that represents the number of siblings of those 10 students. Circle any of the following that could be considered as a variable?

μ $\bar{x}$ s σ

8. In a university, 75% of the students are male and 25% are female. 5% of the male students and 15 % or female own a car. For each statement, determine whether it is true?

a. We can conclude that 20% of the students in the university own a car.

b. We can conclude that the number male students who own a car is equal to the number of female students who own a car.

9. True or False?

a. The sample median is sensitive to outliers.
b. The sample mean is not affected by some observations.
c. The mean and median both describe the center of the distribution.
d. If a density curve has more than one peak, then there is more than one mean.

10. The standard deviation of a distribution describes the degree to which the distribution is

a. bell-shaped
b. symmetric
c. spread (or variable)
d. close to zero

11. Suppose that x is a random variable that follows a normal distribution.
The z-score is found by the computation: $z = (x - \mu)/\sigma$.

a. The z-*score* follows a normal distribution with center 0 and standard deviation 1.
If you only computed $x - \mu$, but forgot to divide it by σ, what kind of distribution does this follow?
What is the center and standard deviation of $x - \mu$?

b. If you did $x/\sigma - \mu$ instead of $(x - \mu)/\sigma$, would you still obtain the same z score? Why or why not?

12. At UNH, each of the heights for women and men follows a normal distribution with the mean for women less than the mean for men. Samuel and Cathy were randomly chosen from the population. It was found that their *z-scores* were .6 for Cathy and .5 for Samuel.

a. Cathy is taller than Samuel.
b. Cathy is shorter than Samuel.
c. Cathy is taller than more women than the number of men Samuel is taller than.
d. Cathy is taller than a larger fraction of women than the fraction of men Samuel is taller than.

Items 13 to 14 refer to the following situation:
Heights of adult women in the U. S. are normally distributed with the population mean of μ = 63.5 inches and the population standard deviation of σ = 2.5. Three medical researchers – Aaron, Brian, and Cathy - are planning to select a random sample of adult women.

Aaron took a sample of size 5.
Brian took a sample of size 50.
Cathy took a sample of size 500 as below.

13. Among $\bar{x}_{Aaron}$, $\bar{x}_{Brian}$ and $\bar{x}_{Cathy}$, what do we know? Please circle the correct answer.

a. It is likely that $\bar{x}_{Aaron} < \bar{x}_{Brian} < \bar{x}_{Cathy}$.

b. It is likely that $\bar{x}_{Aaron} > \bar{x}_{Brian} > \bar{x}_{Cathy}$.

c. It is likely that $\bar{x}_{Aaron} = \bar{x}_{Brian} = \bar{x}_{Cathy}$.

d. It is likely that $\bar{x}_{Aaron}$ is farther from μ than $\bar{x}_{Brian}$ is and $\bar{x}_{Brian}$ is farther from μ than $\bar{x}_{Cathy}$ is.

	Sample Size	Notation for sample mean	Notation for sample standard deviation
Aaron	5	$\bar{x}_{Aaron}$	s_{Aaron}
Brian	50	$\bar{x}_{Brian}$	s_{Brian}
Cathy	500	$\bar{x}_{Cathy}$	s_{Cathy}

14. Among s_{Aaron} , s_{Brian} and s_{Cathy} , what do we know? Please circle the correct answer.

a. $s_{Aaron} = s_{Brian} = s_{Cathy}$

b. $s_{Aaron} < s_{Brian} < s_{Cathy}$

c. $s_{Aaron} > s_{Brian} > s_{Cathy}$

d. It is likely that s_{Aaron} is farther from 0 than s_{Brian} is and s_{Brian} is farther from 0 than s_{Cathy} is.

e. It is likely that s_{Aaron} is farther from σ than s_{Brian} is and s_{Brian} is farther from σ than s_{Cathy} is.

f. It is likely that s_{Aaron} is farther from $\frac{\sigma}{\sqrt{n}}$ than s_{Brian} is and s_{Brian} is farther from $\frac{\sigma}{\sqrt{n}}$ than s_{Cathy} is.

References:

Arcavi, A. (1994). Symbol sense: informal sense-making in formal mathematics, For the Learning of Mathematics 14(3), 24-35.

Eco, U. (1976). A Theory of Semiotics. Bloomington: Indiana University Press.

Font, V., Godino, J., & D'Amore, B. (2007). An onto-semiotic approach to representation in mathematics education. For the learning of mathematics, 27(2), 2-7,14.

Garfield, J. & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. New York, NY: Springer.

Hewitt, D. (1999). Arbitrary and necessary: Part 1 a way of viewing the mathematics curriculum. For the learning of mathematics, 19(3), 2-9.

Hewitt, D. (2001a). Arbitrary and necessary: Part 2 assisting memory. For the learning of mathematics, 21(1), 44-51.

Hewitt, D. (2001b). Arbitrary and necessary: Part 3 educating awareness. For the learning of mathematics, 21(2), 47-59.

Kaplan, J., Fisher, D., & Rogness, N., (2009). Lexical ambiguity in statistics: How students use and define the words: association, average, confidence, random and spread. Journal of statistics education, 18(1). http://www.amstat.org/publications/jse/v18n2/kaplan.pdf

Mayen, S., Diaz, C., Batanero, C. (2009). Students' semiotic conflicts in the concept of median. Statistics education research journal, 8(2), 74-93.

Noth, W. (1995). Handbook of semiotics (Advances in semiotics). Bloomington, IN: Indiana University Press.

Peters, S. Robust understanding of statistical variation. Statistics education research journal, 10(1), 52-88.

Shaughnessy, M. (2007). Research on statistics learning and reasoning. In F.

Watson, J., (2009). The influence of variation and expectation on the developing awareness of distribution. Statistics education research journal, 8(1), 32-61.

Watier, N., Lamontagne, C., & Chartier, S., (2011). What does the mean mean? Journal of statistics education, 19(2). http://www.amstat.org/publications/jse/v19n2/watier.pdf

Zieffler, A., & Garfield, J. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. Statistics education research journal, 8(1), 7-31.

READING COMPREHENSION OF SERIES CONVERGENCE PROOFS IN CALCULUS II

Lisa Mantini, Melissa Mills, and Jesse Johnson
Oklahoma State University

This study is a teaching experiment investigating the effect of reading assignments in Calculus II on student performance. Students from a test section and a control section of Calculus II taught during a summer semester were compared. Both sections used traditional lecture methods, the same on-line homework assignments, and common exams. In one section the students completed additional reading assignments with open-ended questions and in-class quizzes evaluating reading comprehension. The study compares student performance on the common exam covering series convergence and the level of writing fluency in student's written arguments on this exam. In addition, four interviews from comparable students, two in each section, were conducted to investigate the ways in which they read, comprehend, and create a series convergence argument.

Key Words: Infinite Series, Series Convergence Argument, Calculus, Reading Comprehension

Introduction and Literature Review

Reading mathematical arguments is not explicitly taught in the standard Calculus curriculum. Though instructors may lament that their students do not read their textbooks, they do little to help students learn to read except for vague advice like "read with a pencil in hand" (Selden & Selden, 2003). It is commonly acknowledged that many students are unable to use their textbooks as effective educational tools. There could be many reasons why this is so. It may be because of the technical nature of mathematics texts with specialized vocabulary, use of symbols, graphical representations, and condensed syntax (Borasi, Seigel, Fonzi, & Smith, 1998). Some have suggested that "the value of mathematical English to mathematicians is unquestionable, but it may be that this style is not appropriate for textbooks used by students" (Watkins, 1979). Investigating the student-textbook relationship can help us understand the difficulties that students face when reading their textbooks and provide support to improve their mathematical reading skills. Weinberg & Weisner (2011) use reader-oriented theory to analyze textbooks. Their analysis shows that difficulties arise when the author's intended reader, the implied reader (the ideal reader who will be able to understand the material as presented), and the actual empirical reader do not line up.

It has been found that in an inquiry-oriented classroom, reading can serve multiple roles, such as focusing the inquiry, carrying out the inquiry, and communicating results (Siegel, Borasi & Fonzi, 1998). The importance of writing mathematics in Calculus has also been documented (Brandau, 1990; Porter, 1996), however, Porter & Masinglia (2000) have questioned whether the pedagogical value is in the physical process of writing, or just in the mental process of reflection on the mathematics. In either case, requiring students to critically read mathematical arguments and reflect upon their reading is a promising pedagogical technique.

Instructors can play an important role in encouraging students to read mathematics texts in an effective way. Instead of viewing the text as merely a source of information, instructors can encourage reader-centered approaches by using the text as a topic of discussion (Weinberg &

Weisner, 2011). The Reading to Learn Mathematics project incorporates engaging students in a wide array of mathematics-related texts including both technical and expository texts. These reading and discussion activities were used in high school classrooms to encourage students to develop a deeper understanding of the content (Borasi & Siegel, 1990; Borasi, et al., 1998; Siegel, Borasi, & Fonzi, 1998). In an introductory calculus course at the university level, Stickles & Stickles (2008) used reading assignments to encourage students to read their text before class. The worksheets consisted of questions to answer and blanks for them to fill in as they read the material. Questions focused on main ideas, notation, vocabulary, and examples presented in the text. A more student-centered way to assess student reading is the use of reading questions. In an undergraduate physics course, students were required to pose questions about their pre-assigned reading via email before coming to class. The questions were graded on a scale from 0-4, assessing the depth of understanding demonstrated by the question. The instructor either answered the questions individually by email, or used the questions in class to motivate discussions (Henderson & Rosenthal, 2006).

We believe that there is value in encouraging students to read mathematical English. This is especially true at the level of Calculus II because many of these students will be continuing their study of mathematics, science or engineering where their ability to communicate mathematical ideas and read mathematical texts will be crucial to their success. In particular, we chose the content area of sequences and series arguments because the students have to give written justification for convergence or divergence that goes beyond mere computation. This may be the first time in the students' mathematical experience when they need to write a justification for their argument. Reading such arguments in the standard mathematical style is a skill that students must learn if they are to produce such arguments on their own. The proof-like structure of these solutions may be difficult for students to comprehend, and we hypothesize that encouraging students to read such arguments will help them to write their own.

The literature shows that students think about series in a wide variety of ways, including visual, verbal and algebraic, shaped by their own view of their role as a learner (Alcock & Simpson, 2004; Alcock & Simpson, 2005). A number of non-traditional methods for presenting the idea of sequence convergence have been proposed using activities to help students make sense of the formal language (Burn, 2005; Roh, 2008; Roh, 2010).

This study will address the need to improve student success in understanding and constructing series convergence arguments by adding some assignments and quizzes that directly assess their reading comprehension of these types of solutions. We created a series of "comprehension quizzes" designed to be suitable for use as supplements to a traditionally taught Calculus II course. Each quiz begins with an argument that could serve as a model for proofs that the students were expected to construct for themselves. In fact, many of the arguments that were used in this study were adapted from or excerpts from Stewart's Calculus with Early Transcendentals (Stewart, 2008).

Each argument was followed by a number of comprehension questions based on the model designed by Mejia-Ramos, Fuller, Weber, Rhoads, & Samkoff (2012). This model consists of seven dimensions:

Local Assessments:

1. Meanings of terms and statements
2. Logical status of statements and proof framework
3. Justification of claims

Holistic assessments:

4. Summarizing via high-level ideas
5. Identifying the Modular Structure
6. Transferring the general ideas to another context (writing another proof)
7. Illustrating with Examples (how it relates to specific examples)

We wrote assessment questions related to six of the seven assessments, leaving out only item (2.) on the logical status of statements and the proof framework. This item seemed to involve a higher level of sophistication in student thinking and higher ability to think abstractly than what we typically expect in Calculus II, so it seemed more difficult to incorporate this assessment into activities in Calculus II. We feel that the other assessments worked well in Calculus II. The reading comprehension quizzes are not included in this paper, but the type of questions asked are very similar to our Interview questions, which are included in Appendix B.

Research Questions

1. Will students' facility in determining series convergence or divergence improve after activities that emphasize and assess reading comprehension of series arguments?
2. Will students have more fluency in writing justifications of series convergence or divergence after assessments of reading comprehension of convergence arguments?
3. Do students read mathematical arguments differently or comprehend more of what they read after activities that emphasize and assess reading comprehension?

Methods

Two sections of Calculus II were taught during a summer semester by instructors with similar styles and similar teaching experience. At this institution, Calculus II includes techniques of integration and applications, sequences and series, parametric curves, and polar coordinates. Both instructors were advanced doctoral students in mathematics who had not previously taught Calculus II but who had positive experiences teaching Calculus I. Both sections were taught in a traditional way, with the majority of each class period devoted to lecture and additional time spent on class discussion and problem solving by students. Students self-selected between the two sections, which met at the same time, with 19 students enrolling in the first section (control) and 29 students enrolling in the second (test). The two sections used identical examinations given four times during the semester and identical assignments in an online homework system. After the first examination, students in the test section participated in additional in-class activities which emphasized comprehension of mathematical passages read by the students. They also completed several quizzes assessing reading comprehension of series convergence arguments taken from their textbook (Stewart, 2008).

Data Collection

The first in-class examination, covering techniques of integration and applications, was used to compare the level of the students in both sections. After names were removed and codes were assigned, copies of these exams were distributed to the three researchers.

The second in-class examination, covering convergence of series of constants, was used to answer the first two research questions. After names were removed and codes assigned, copies were distributed to the three researchers. The exam included five problems in which the students were to determine the convergence or divergence of a given series and write their justification. It also contained one problem that used the integral test estimation theorem, and three conceptual true or false questions. There were no questions on the second exam that directly addressed their reading comprehension abilities.

Student volunteers were recruited between the first and second exams. After the final exam, we interviewed two students from each of the classes from among the volunteers. The students were chosen to have a grade of C or better on the first midterm. The scores of the students from the control section were an A and a B, and both students from the test section scored a C on Exam 1. Each student was interviewed separately by a member of the research team shortly after their final exam for the class. The interview questions were similar in style to the questions on the reading comprehension quizzes that the students completed in the test section. The interview subjects were presented with an example solution in which an interval of convergence for a power series was calculated. First, the students were asked to summarize the solution and then give the purpose of each paragraph. The students were also presented with four questions related to the solution, and asked to think aloud as they answered the questions. Then they were then given a similar power series and asked to find the interval of convergence. When the students completed the problem, they were asked about different aspects of the course and what contributed to their ability to read and comprehend series convergence arguments. The interview questions are included in Appendix B.

Data Analysis

To determine the effectiveness of reading comprehension activities on student performance, we first analyzed their performance on two common exams (Porter & Maslingila, 2000), one exam occurring before the reading comprehension activities, and one after. This portion of the analysis is used to address the first research question.

Two researchers independently scored Exam 1 for all students, using their own rubrics. Scores were averaged to obtain a combined score for each student. The distribution of scores from each section was tallied and compared.

Two researchers independently scored Exam 2 for all students, using a common rubric. Scores were averaged to obtain a combined score for each student. The distribution of scores from each section was tallied and compared. We also compared the net change in performance from Exam 1 to Exam 2 for students from each section. In addition, we counted the number of students who gave a correct determination of convergence or divergence, with a correct reason quoted, for each of five series given on Exam 2 and compared by section.

In order to answer the second research question, the researchers assessed the fluency of the arguments written by students on Exam 2 justifying their conclusion of convergence or divergence for five different infinite series. We developed a fluency scoring rubric which assigned a numerical score of 0, 1, or 2 points to three parts of each argument: the introduction, the body of the argument, and the conclusion. This score reflects the readability of their written argument, including their organization, use of proper notation, and expression of ideas in complete sentences. The researchers jointly assigned a numerical score of 0 through 5 to each written argument. We did not use a score of 6, even though the rubric would have allowed a maximum score of 6, since it would have occurred extremely rarely.

In order to answer the third research question, the interviews were transcribed and two of the researchers read the transcripts looking for general themes in the discussions and differences between the students in the two classes. The researchers made notes about the different ways the students were interacting with the text. By comparing notes and selected excerpts, the two researchers agreed upon four themes: ability to summarize the solution, mathematical accuracy, attitudes towards the written solution, and confidence in their answers.

To demonstrate the ability to summarize the solution, we wanted the students to make it clear that they had a coherent overall picture of the solution. We were looking for the students to be able to use their own words to talk about the solution, and the ability to do so without prompting. In particular, we wanted them to go through the solution in order and to not be bogged down in the details of the solution, but to see the bigger picture.

When analyzing the mathematical accuracy of student responses, we were looking for how closely the students' solutions aligned with what was expected as a correct solution. We were also looking for the students to demonstrate evidence of their understanding of the material in the solution.

We then looked at the students' attitudes toward the written text and the exercise of reading a worked solution. We looked for whether the students appreciated the text, read it carefully, referenced the text when answering questions, or made any comments about how they felt about the text.

Throughout the interview the students made comments that reflected their confidence in both comprehending the written argument and producing their own solutions. They also made comments about their confidence in the material from the entire course.

Once the researchers had established these themes, they read the transcripts again, making comments and choosing excerpts that exemplified the themes in each interview. The results are organized by four themes and summarized in the results section.

Results

Student Performance on Exams

Exam 1 was scored individually by two researchers with individual rubrics, and each student was assigned the average of these two numbers as their Exam 1 score. The two researchers' scores agreed fairly closely, with an average score difference of 3.86 points and a standard deviation of 3.3 points. Scores assigned by the two researchers differed by 0-13 points, but only four scores differed by more than 9 points. It may be worth noting that the average difference was twice as large (5.1 points) on students in the bottom half of the population (scores below 65) as it was on the top half of the population (2.6 points).

Based on the score distribution, the Exam 1 scores were sorted into five groups, as suggested by the data: the A group, with scores from 90-105, the BC group, with scores from 75-89, the CD group, with scores from 60-74, the High F group, with scores from 45-59, and the Low F group, with scores below 44. The number of students with scores in each group and the section averages are shown in Table 1.

Table 1
Exam 1 Scores

Section	0-44	45-59	60-74	75-89	90-105	Average
Control	5	3	3	3	3	61.2
Test	6	5	7	5	4	65.4

Exam 2 was scored individually by two researchers with a common rubric, and each student was assigned the average of these two numbers as their Exam 2 score. The two researchers' scores agreed fairly closely, though agreement was *less* with a common rubric than with individual rubrics as on Exam 1. The average score difference was 4.07 points with a standard deviation of 3.66 points. Scores assigned by the two researchers differed by 0-17 points, but only three scores differed by more than 9 points. It was still the case that the researcher's score difference was greater on the bottom half of the class than on the top half, but not by quite as much as for Exam 1. For the bottom half of the class, with scores less than 68, the researcher's scores differed by an average of 4.95 points, as compared to an average 3.18 point difference on the top half.

The Exam 2 scores were sorted into the same five groups as for Exam 1, which continued to reflect the data distribution. The number of papers scoring in each group on Exam 2 and the section average scores are shown in Table 2.

Table 2
Exam 2 Scores

Section	0-44	45-59	60-74	75-89	90-105	Average
Control	4	1	3	7	2	64.8
Test	6	8	1	4	8	65.9

The researchers compared the Exam 2 scores with Exam 1 scores, in order to try to determine any possible effect on score improvement that the reading comprehension activities might have had. It appears that mid-range and low-scoring students improved more in the control section than in the test section: there were fewer F's on Exam 2 than Exam 1 in the control section and more B's. However, students performing well on Exam 1 fared better on Exam 2 in the test section than in the control section: all students with A's on Exam 1 maintained their A on Exam 2 in the test section, whereas all A students on Exam 1 in the control section dropped to the BC group on Exam 2. Several students from lower groups also improved significantly and earned A's on Exam 2 in the test section. Table 3 shows the relative change from Exam 1 to Exam 2 in each section, sorted by the initial Exam 1 score grouping.

Table 3
Net Change in Score from Exam 1 to Exam 2

Section	0-44	45-59	60-74	75-89	90-105	Average
Control	+9.6	+9.7	+10.2	-1.7	-13.5	+3.6
Test	+11.6	+2.4	-6.9	-2.2	-4.5	+0.5

Exam 2 contained five problems which asked students to correctly determine the convergence or divergence of a series, with justification. The questions from Exam 2 are reproduced in Appendix A. For these first five questions, the researchers counted the percentage of the students in each section who answered the problem correctly, that is, who correctly determined the convergence or divergence of the series based on a justification that was essentially correct. The correctness percentages are summarized in Table 4.

Table 4
Percentage of Correct Solutions for Exam 2 Problems 1-5

Section	1	2	3	4	5
Control	76%	**65%**	65%	29%	**59%**
Test	**85%**	44%	63%	**41%**	52%

This analysis showed that two of the five questions were more likely to be answered correctly by students in the test section, two of the five were more likely to be answered correctly by students in the control section, and one problem was equally likely to be answered correctly by students in either section. This indicates that the reading comprehension activities completed by students in the test section did not appear to give them an advantage in correctly determining the convergence or divergence of a series. The bold numbers indicate the group of students with a higher likelihood of answering the given problem correctly.

Writing Fluency

Problems 1-5 on Exam 2 asked the students to determine whether a given infinite series converged or diverged, and to write a justification of their answer. We wished to determine if students who had completed the reading comprehension activities would be able to write more fluent justifications of their conclusions regarding series convergence or divergence. To this end, we subdivided student arguments concerning series convergence or divergence into three parts: the Introduction contains the identification of the type of series or identifies a relevant convergence test; the Body of the argument includes the justification that a particular test is applicable or contains the steps needed to implement the test; and the Conclusion states the determination of either convergence or divergence for the series in question based on the interpretation of the test that was used.

We developed a Fluency Rubric which assesses the fluency of a student's argument by assigning a score of 0, 1, or 2 points to each of the three parts of the argument. The Fluency Rubric is summarized in Table 5. This rubric would lead to possible fluency scores of from 0 to 6 for a given student's argument; however, of 220 arguments scored, only 1 or 2 would have merited a score of 6. So we opted to use 5 as the maximum possible score for papers scoring either a 5 or 6 by this rubric.

Table 5
Fluency Rubric

Introduction	0	No evidence of an introduction
	1	Attempt at introduction: naming test, phrase or sentence fragment
	2	Correct introduction in complete sentence or phrase
Body	0	No justification or no relevant work
	1	Attempt made to justify or apply test criteria
	2	Checking and applying all test criteria accurately, laying out argument clearly
Conclusion	0	No conclusion
	1	Some conclusion shown in a word or phrase
	2	Correct conclusion in a full sentence or phrase

Tables 6 and 7 summarize the results of the fluency scores for students from the control and test sections. The tables indicate the average fluency score in that section for each problem, along with the percentage of students in that section scoring a 4 or 5, indicating a high level of fluency in the argument, and the percentage of students in that section scoring a 0 or 1, indicating a very low level of fluency in the written argument.

Table 6
Fluency Scores for the Control Section on Exam 2 Problems 1-5

	1	2	3	4	5
Section average	2.76	2.47	2.94	2.88	1.82
Percentage 4-5	29.4%	23.5%	11.8%	29.4%	5.9%
Percentage 0-1	17.6%	29.4%	0.0%	11.8%	41.2%

Table 7
Fluency Scores for the Test Section on Exam 2 Problems 1-5

	1	2	3	4	5
Section average	3.44	2.44	3.07	3.26	2.30
Percentage 4-5	40.7%	37.0%	40.7%	44.4%	25.9%
Percentage 0-1	3.7%	37.0%	22.2%	11.1%	33.3%

Notice that proportionately more students write highly fluent arguments in the test section than in the control section on every single problem. This would support a conclusion that reading comprehension activities lead to greater writing fluency on series convergence arguments. But Tables 6 and 7 also show that on two of the problems, proportionately more students in the test section write the least fluent arguments. This might support a conclusion that the reading comprehension activities are most helpful for better students and are less likely to be helpful for students who are struggling.

Interview Data

When presenting the results of the interview data, we will use pseudonyms for the participants. The participants from the control section will be called Alan and Andrew, and the participants from the test section will be called Blake and Brian. While we cannot make any overall claims because of the small sample size, the interviews showed marked differences between the two classes in the four areas selected for analysis. This supports a conclusion that the comprehension quizzes affected the way that these two students approached reading the solution.

Ability to summarize the solution. Alan and Andrew both needed help from the interviewer to come up with a summary for the written solution. Alan asked the interviewer what was meant by "please explain the solution," then gave a rather garbled explanation.

> … well, first the ratio test, clearly, and then they canceled out, and you come to the radius of convergence where, uh, minus, the absolute value of minus three is less than 1. So, you come to this part where two is less than, which is less than four. And then later you check that just to make sure of whether, for the interval

of convergence, … whether the 2 and 4 would be included.

Andrew also needed prompting, asking “Oh, I need to read it?” before giving a similarly garbled summary. When asked to clarify something from his explanation, Andrew says “... Hmm… throw me a curve ball here...”.

Blake and Brian, on the other hand, were able to accurately summarize the solution without prompting. Blake gave a very concise initial summary:

> … to find the radius of convergence you have to do the ratio test, and you equate that to less than one. And that will tell you your radius of convergence and also the interval, but it’s not conclusive for the endpoints, so you have to do a different test for each endpoint.

When asked to explain what happens in each paragraph, Blake’s explanation was similarly concise and accurate. Brian’s explanation is longer and much less eloquent, but still accurate, and did not require further prompting.

Mathematical accuracy. The most difficult question for all four students turned out to be finding the domain of the power series in the solution. Brian came the closest to the accurate solution when he guessed “I want to say it has something to do with the interval of convergence...” but he was still unable to make the connection without help from the interviewer. The other students had no response for this question. It is possible that the equivalence of the interval of convergence with the domain of the power series function was not emphasized in class. All four students displayed a misunderstanding of the implications of divergence. Blake, when describing what would happen if a number outside of the interval of convergence was substituted for x, said,

> “we won’t get an exact value for the function. It will, like, go off to infinity, or else… basically it diverges away from any certain value, so it, but I guess the reason that I was thinking that it would still be in the domain, just not the interval is because you could still estimate it. You could come up with an answer, it just wouldn’t be an exact answer, so technically, I guess it’s outside the domain.”

These may be common misconceptions for students in Calculus II.

When asked to find the interval of convergence for a similar power series, both students from the control section failed to check the endpoints even though they had just discussed this part of the written solution. Andrew computed the interval correctly, but did not even acknowledge the need to check the endpoints. Alan also correctly calculated the interval, but attempted to use the Ratio Test to check the endpoints, and gave up when he kept getting 1 for the limit. Both Blake and Brian were aware of the need to check the endpoints and did so accurately. They also were very fluent in explaining their solutions.

Attitudes towards the written solution. Both students from the control section appeared dubious about reading the solution that was presented to them. After giving his summary of the solution, Alan admitted that he didn't read it very carefully. He also admitted that he is “not particularly confident” in reading this type of solution and that the words are

> “not that helpful... if they just had, like, say limit comparison test, right next to this stuff, like, just state the test that they use and show the work in-between.

> That would be far more useful than just, than having all this stuff written out."

Andrew, similarly, didn't seem to completely understand the solution. For example, at one point he said

> "I don't see why that gives cause for using the limit comparison test. I don't think the limit comparison test can be very helpful."

He later said, "I feel like this is a mind game." When asked to elaborate on his frustrations with the material, he said "This is something that I have to beat into my head." and later predicted that most students in the class pushed the "I believe" button and just memorized the different types of solutions.

The students from the test section seemed much more comfortable with reading the solution. As noted above, they both gave accurate descriptions of the solution, suggesting that they had read it reasonably carefully. Unlike the students from the control section, they did not complain about parts of the solution they didn't understand, and their accuracy in producing their own solution to a similar problem indicates that they read and understood the given solution.

Blake explicitly said that the comprehension quizzes in class helped him:

> "I thought the little quizzes helped, too. Because it's like, I don't know. Personally, I remember things better whenever I get it wrong on something, so whenever I get it wrong then I can see the mistake and learn more from it than I will getting it right. So, the quizzes actually helped… Especially the ones where they were set up more like this and they kinda gave it out and you could look over this and answer questions to it… Whenever we had a quiz that had an example worked out and then we used the example to figure it out, it actually helped a lot for me to understand it."

Brian gave more insight into the way that he thinks about reading mathematical arguments.

> Math is a language of its own. It's just like any language. When you first start reading it it's hard to read. And so it, I guess it takes some training, and, uh, a definite amount of attention and understanding of what sentence, or what formulas you read beforehand, and how they interact with each other. Um, I do have a lot of difficulty reading math, and my, obviously by answering the problems, I'm familiar with the formulas, so… like, reading a story, you don't know what's happening next, but when I'm reading this, because I know what's going to happen in the next chapter, or you know, in the next page, or whatever. I know that Dumbledore dies, basically. And, uh, it, um, I know what to expect… and it's structured in the manner of my thought process. You get the interval, and then you test both endpoints. And it did do left end first, right end second, so it's very familiar to my thought process, and so I was able to read it because that's the way my mind would work.

Notice that Brian considers the types of arguments in the written solution quite natural. Contrasting this with the attitudes of Alan and Andrew, who expressed the opposite view, we suspect that Brian's sentiments were influenced by the comprehension assignments and quizzes.

Confidence in their answers. There was a remarkable difference in the appearance of confidence between the students from the different sections. Alan continually qualified his answers with remarks such as "unless I'm mistaken" and twice during the interview, asked "Have I messed up yet?" He also told the interviewer "I was never very good at that part" when asked about finding the domain and claimed not to remember a number of topics.

Andrew told the interviewer that a number of concepts were not explained in class and complained that the sigma notation looked "weird", "strange", or "confusing". In particular, Andrew remarked:

> I appreciate what's happening here, but I don't like doing it because it's just different. This kind of math looks different from any other kind of math I've ever had to do.

Blake and Brian, on the other hand, never qualified their answers during the interview. Even though they were equally lost on the domain question, they reasoned their way through it without becoming frustrated. Brian noted that he lost points on an exam for leaving out the limit sign on his solution, but stated that he understood the problem anyway. So, the students in the test section were more able to work with the mathematics with confidence even when they acknowledged that they had made an error.

In summary, the students in the test section did appear to read and interact with the written solution in a different way. They were better able to summarize the solution without prompting, they were better able to use the structure of the solution to guide the construction of their own solution, they were more positive about the usefulness of the written solution, and they were overall more confident with the content. It should be noted that the two students in the test section both made C's on Exam 1, but they both pulled their grade up to an A by the end of the semester. In contrast, the two students in the control section had earned grades of A and B on Exam 1, but both dropped one letter grade by the end of the semester, finishing with grades of B and C, respectively.

Discussion

Reading comprehension activities do seem to show some positive effects on aspects of student performance in Calculus II, but these effects seem to be concentrated in certain areas of student performance more than others.

First we consider a student's ability to correctly determine convergence or divergence of an infinite series. We assessed this ability in two ways: both by comparing student scores on Exam 2 in the test and control sections, and also by comparing the overall proportion of students in each section who made the correct determination of convergence or divergence on the five most routine problems on Exam 2. In comparing students' total scores on Exam 2, no clear advantage seemed to accrue to students in the test section over those in the control section. It is possible that the best students in the test section were able to obtain an advantage over other students after reading comprehension activities, but this is difficult to claim with any certainty. We did not see an advantage for students in the test section over those in the control section with regard to determining convergence or divergence of an infinite series.

With regard to fluency of their written arguments, it does appear that reading comprehension activities led to better fluency in written arguments for students in the test section over those in the control section. Having students read and answer questions about well-written arguments seems to help students to know what a well-written argument is, and so helps them to write more fluent arguments themselves.

Finally, our interview data seem to show several positive effects of reading comprehension activities on how students read an argument, on how they think about what they have read, and on how they explain the argument to a researcher. Students who had undergone training in reading comprehension showed a better ability to summarize an argument they read, and they were more accurate in answering certain types of questions correctly. Students who did not undergo training in reading comprehension seemed to express frustration about the mathematics and seemed to lack confidence in their ability to answer questions. On the other hand, students who had undergone reading comprehension activities were confident in their ability to answer questions and expressed a positive attitude about the course and about the mathematics.

Limitations of this study would include the small sample size, particularly in the number of interviews conducted. Our four interview subjects may or may not be representative of the student population in this study. We should probably note that the study was conducted during the summer semester. The audiences for summer school classes at this University may include a higher proportion of students who are repeating a class for the second time than is typical during a spring or fall semester, and so it is likely that a higher proportion of students struggle with this material in the summer semester than in spring or fall semesters.

Despite these limitations, we believe that reading comprehension activities show promise in helping our students to achieve greater writing fluency and a greater level of confidence in their understanding of the material. We believe that it would be worthwhile to investigate the value of reading comprehension activities throughout the calculus sequence.

References

Alcock, L. & Simpson, A. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics*, 57(1), 1-32.

Alcock, L. & Simpson, A. (2005). Convergence of sequences and series 2: Interactions between nonvisual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics,* 1, 77-100.

Borasi, R. & Siegel, M. (1990). Reading to learn mathematics: New connections, new questions, new challenges. *For the Learning of Mathematics 10*(3), 9-16.

Borasi, R., Siegel, M., Fonzi, J, & Smith, C. (1998). Using transactional reading strategies to support sense-making and discussion in mathematics classrooms: An exploratory study. *Journal for Research in Mathematics Education, 29*(3), 275-305.

Brandau, L. (1990). Rewriting our stories of mathematics. In A. Sterrett (Ed.) *Using Writing to Teach Mathematics* (pp. 73-77*).* Washington, D.C.: Mathematical Association of America.

Burn, B. (2005). The vice: Some historically inspired and proof-generated steps to limits of sequences. *Educational Studies in Mathematics,* 60, 269-295.

Henderson, C. & Rosenthal, A. (2006). Reading questions: Encouraging students to read the text before coming to class. *Journal of College Science Teaching, 35*(7), 47-50.

Mejia-Ramos, J., 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.

Porter, M.K. (1996). The effects of writing to learn mathematics on conceptual understanding and procedural ability in introductory college calculus. *Dissertation Abstracts International, 58*, 2576A.
Porter, M. K. & Masingila, J. O. (2000). Examining the effects of writing on conceptual and procedural knowledge in calculus. *Educational Studies in Mathematics,* 42, 165-177.
Roh, K. H. (2008). Students' images and their understanding of definitions of limit of a sequence. *Educational Studies in Mathematics,* 69, 217-233.
Roh, K. H. (2010). An empirical study of students' understanding of a logical structure in the definition of limit via the ε-strip activity. *Educational Studies in Mathematics,* 73, 263-279.
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*, 4-36.
Siegel,M., Borasi,R. & Fonzi, J. (1998). Supporting students' mathematical inquiries through reading. *Journal for Research in Mathematics Education,* 29(4), 378-413.
Stewart, J. (2008). *Calculus: Early Transcendentals* (6th ed.). Mason, OH: Cengage Learning.
Stickles, P. R & Stickles, J. A., Jr. (2008). Using reading assignments in teaching calculus. *Mathematics and Computer Education,* 42(1), 6-10.
Watkins, A. E. (1979). The symbols and grammatical structures of mathematical english and the reading comprehension of college students. *Journal for Research in Mathematics Education*, 10 (3) 216-218.
Weinburg, A. & Wiesner, E. (2011). Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics,* 76, 49–63.

Appendix A: Exam 2

1. Determine whether the series converges or diverges. State which test(s) you use and show your work.
2. Determine whether the series converges or diverges. State which test(s) you use and show your work.
3. Determine whether the series converges or diverges. State which test(s) you use and show your work.
4. Determine whether the series is absolutely convergent, conditionally convergent, or divergent. State which test(s) you use and show your work.
5. Determine whether the series converges or diverges. If it converges, find the sum.
6. Do the following for the series .
   a. Evaluate .

b. How many terms of would we need to find its sum to within 0.001?

7. Short answer questions.

   a. True or false: If diverges then also diverges. Explain.

   b. Give one example of a series so that $\lim_{n\to\infty} a_n = 0$ and converges. Explain why your series converges.

   c. Give one example of a series so that $\lim_{n\to\infty} a_n = 0$ and diverges. Explain why your series diverges.

Appendix B: Interview Questions

Review the exercise and solution carefully, and answer the questions that follow.
Exercise: Find the radius of convergence and interval of convergence for the

series .

Solution: Since $a_n = (-1)^n \frac{(x-3)^n}{(2n+1)}$, we can apply the Ratio Test to find

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{(x-3)^{n+1}}{(2n+3)}\frac{(2n+1)}{(x-3)^n}\right| = |x-3|\lim_{n\to\infty}\frac{2n+1}{2n+3} = |x-3|.$$

The Ratio Test says that the series converges when $|x-3|<1$. Thus, the radius of convergence is $R=1$.

Now we will find the interval of convergence. The Ratio Test told us that the power series converges when $|x-3|<1$, which implies that $2<x<4$. The Ratio Test gives us no information when $|x-3|=1$, so we much check those two cases.

When $x=2,$ the series becomes . This series can be compared with , which is a constant times the harmonic series. Since $\frac{1}{2n+1}<\frac{1}{2n}$, we can't do a direct Comparison Test. We need to use the Limit Comparison Test.

$$\lim_{n\to\infty}\left|\frac{a_n}{b_n}\right| = \lim_{n\to\infty}\left|\frac{2n}{2n+1}\right| = 1.$$

By the Limit Comparison Test, since diverges, so does the given power series when $x=2$.

When $x = 4$, the series becomes . The signs strictly alternate, (i) $\frac{1}{2(n+1)+1} < \frac{1}{2n+1}$, and (ii) $\lim_{n\to\infty} \frac{1}{2n+1} = 0$. So the power series converges by the Alternating Series Test when $x = 4$.

Thus, the interval of convergence is $I = (2,4]$.

Questions:

1. Please explain the solution in your own words.
2. Explain the purpose of each paragraph in the argument.
3. Write out the first three terms of the power series.
4. What is the domain of the function $f(x) =$?
5. Is the sum convergent or divergent for $x = 5$? Explain.
6. Does the series $f(x) =$ converge or diverge? Explain your answer.
7. Find the radius of convergence and interval of convergence for the power series $\sum_{n=1}^{\infty} (-1)^n \frac{(x-1)^n}{3^n (n^2)}$.
8. Was there anything that you did in class or on the homework that helped you to answer these questions?

TO REJECT OR NOT REJECT: ONE STUDENT'S NON-NORMATIVE DECISION PROCEDURE FOR TESTING A NULL HYPOTHESIS

Michael McAllister
Arizona State University

The purpose of this study was to gain insight into how exposure to hands-on and computer resampling methods affected a statistically naïve student's emergent understandings of statistical inference. In this study, simulation design activities provided a vehicle for engaging a student with the core ideas of hypothesis testing. The results highlight challenges the student experienced in coordinating the components of the logic into a coherent scheme of ideas and sheds light on aspects of engagement which need to be emphasized in order to resolve the inherent conceptual difficulties associated with reasoning that invokes a modus tollens-like argument. Moreover, I report on a heuristic the student used to make his inferential decisions—one that does not produce correct inferences. I've termed this the "similarity heuristic" because of a specific similarity relationship the student would look for and then use as a method for rejecting or not rejecting the hypothesis being tested.

Key words: Statistical inference, Statistical reasoning, Hypothesis testing, Resampling, Simulation

Introduction and Literature Review

Statistical inference is arguably one of the most important ideas we might expect students to understand. The ability to draw inferences from data has an enormous impact on society and is critical to the advancement of knowledge. Inference is now applied in a wide range of scientific disciplines and given its extraordinary range of applications the question of how to support the development of a coherent understanding of statistical inference has taken on an increased importance. The traditional approach to teaching statistical inference is through probability based distributions couched in abstract theory and formal language. Statisticians of generations past invented these analytical methods in part because direct simulation through resampling was simply too slow to be practical (Cobb, 2007). R. A. Fisher in the 1930's recognized the usefulness of generating an empirically created sampling distribution, under the assumption of a null hypothesis, but didn't have the computing power to rely on this direct approach (Yu, 2007). Today, however, the computing power to rely on direct simulation methods is easily accessible and this has created a growing movement concerned with how we teach statistical concepts. Many educators (e.g., Chance (2006), Erickson (2006), Cobb (2007), Rossman (2008) and Garfield & Ben-Zvi (2008)) now regard parametric methods as too formal an introduction for most students and they advocate a new equilibrium that opens the door for computer simulation activities as a way to help students understand the difficult concepts which underlie how statistical decisions are made.

Computers have already had a tremendous impact on the practice of statistics at the tertiary level. In addition to statistical packages such as SPSS which carry out the tedious calculations required in estimation and hypothesis testing procedures, there is a growing assortment of interactive simulation software which enables the user to mimic the real-life repeated random sampling which gives rise to such things as the Central Limit Theorem, confidence intervals, and hypothesis tests. Rossman (2008) argues that simulations put more emphasis on the core ideas enabling the learner to experience firsthand how a statistic of interest varies from sample to sample, how an empirical sampling distribution evolves with an increasing number of resamples, and the important role played by sample size.

Mills (2002) provides an overview of the literature on the use of computer simulation methods from 1983 to 2000. In these reviewed articles many researchers in statistics education recommended the use of computer simulation methods to teach abstract concepts in statistics. One suggested advantage was simulation's ability to utilize the power of concrete illustration to ease logical difficulties and enhance understanding. The consensus was that computer simulations are arguably instructionally productive. Only a tiny subset of these studies, however, collected empirical data and the ones that did relied heavily, if not exclusively, on quantitative measures to determine the effect of computer simulation methods on student achievement. One common approach was the comparison of pre and post-test performance after exposure to an intervention. For example, delmas, Garfield and Chance (1999) demonstrated a powerful effect of using computer simulation on students reasoning about sampling distributions and the Central Limit Theorem. Several more recent publications have also suggested that improved instructional results can be achieved by using good simulation tools and activities (Lipson, 2002; Chance, delmas & Garfield, 2004). The quantitative results of these studies indicated that "something" important happened between the pre and post-test measures but without additional evidence it's not possible to reveal exactly what that was.

Saldanha (2004) reported on a series of classroom teaching experiments that engaged high school students with instructional tasks in which they designed the components of the simulation in the context of modeling contextual scenarios involving hypothesis tests. The intent was to use computer simulations and the interactions that flowed out of that engagement to help students understand the vital connections between sample, population, and the sampling distribution on which an inference is based. The study is notable because it is one of the few to actually characterize the reasoning that emerged as students engaged in instructional activities centered on the use of computer simulations. As part of this larger study (Saldanha 2004), Saldanha & Thompson (2007) continue the discussion by reporting on key developments and critical shifts that unfolded over a series of 3 consecutive lessons as students engaged in both concrete and computer simulated sampling activities. The report folds back from the data to characterize how instruction shaped the students' conceptions of sampling distributions and the inferences that can be made based on these collections. More recently, Saldanha (2011) reports on a single simulation activity in which a group of high school students encountered severe conceptual difficulties as they grappled hard with the crucial process of turning a phenomenon of interest into first a statistical question and then into a stochastic experiment in order to judge whether a particular event was unusual.

Educators (e.g. Chance & Rossman 2006; Cobb 2007; Kaplan 2007; Rossman 2008; Garfield & Ben-Zvi 2008) have recently put forth altogether different approaches to teaching statistical inference in an introductory college course. A distinctive feature of these curriculums centers on their use of resampling methods--both hands-on and computer simulated--as the entry point for developing students' inferential concepts. Inference is introduced devoid of the mathematical computations and formulas so that students can focus on what the fundamental ideas of null hypothesis, distribution and p-value mean. Cobb (2007) suggests placing inference as the focal point of a course that introduces statistical inference by three R's: (a) Randomize the data production; (b) Repeat by a computer simulated model to see what is typical; and (c) Reject any null hypothesis model that puts your observed sample in the tail of the empirical distribution.

Research Questions

The ideas and suggestions of Cobb (2007), Rossman (2008) and Garfield & Ben-Zvi (2008) merit serious attention but there is as of yet little empirical research that documents how the impact of these methods actually shape students' reasoning and understanding. While numerous studies have now used computer simulation as a vehicle to help students build the requisite imagery underlying inference, they haven't necessarily characterized the ways of thinking that students express as they choose procedures and work their way from claims to conclusions. My review of the literature suggests a need for studies that shed light on students' cognitive processes as they design and enact simulation activities. Hence, the goal of this study was to help fill this gap by exploring the development of a single student's thinking in relation to his engagement in specially designed activities involving resampling simulations. The activities, aimed to engage the student with the logic of hypothesis testing through the creation of simulation models that resemble the phenomena to be investigated. Questions of interest included: (a) what ways of thinking--interpretations, understandings and imagery--express themselves as the student engages in the instructional activities? (b) What conceptual difficulties did the student encounter? (c) What aspects of engagement in these activities hindered or moved his thinking forward in productive ways?

Theoretical Perspectives

Four basic theoretical perspectives underlay this study and were drawn on in the design of the instructional activities and the data collection and analysis.

I first drew upon constructivism as elaborated by von Glasersfeld (1995) as a way to understand learning and learners. Adopting this perspective had several implications. First, while I don't believe that telling is an anti-constructivist pedagogical action, a constructivist perspective urges one to keep in mind that regardless of how clearly a concept is explained the student will construct his own meaning for it and this meaning may be dramatically different than your own. Second, it forces the researcher to constantly question his interpretations of how the student understands what he (the researcher) takes as normative reasoning. Since we're trying to affect the student's understanding of situations in ways that support the emergence of particular types of reasoning a constructivist perspective requires the researcher to constantly be thinking of ways the student might be thinking in order to adjust his actions accordingly. Third, it is useful as a way to describe ways of knowing particular ideas that operationalize what it is a student should understand in order to know an idea in a particular way. Fourth, von Glasersfeld describes an analytical method called conceptual analysis for creating a hypothetical conceptual system that approximates the student's own set of conceptual operations that he uses to know something in the way he apparently does. Conducting conceptual analysis entails imagining what *something* the student has in mind in the context of his discussing that *something*. In other words, the researcher puts himself into the position of the student and attempts to imagine the mental operations that he would need or the constraints he would have to operate under to behave as the student did.

The second perspective was Thompson's (1994) theory of quantitative reasoning. Thompson's (1994) theory of quantitative reasoning is about people conceiving situations in terms of quantities. It describes the kind of reasoning I intended for the student in this study to develop with respect to measuring and quantifying unusualness. Key to quantitative reasoning as characterized by Thompson is the idea that a coherent conception of a quantity entails conceptualizing situations in ways that support apprehending the attributes of interest embedded within them as measurable. Unusualness is one of the key attributes embedded in the logic of hypothesis testing and intuitive ideas about how unusual a sampling event is are not enough; that is, a "gut feeling" that a sampling outcome is unusual is an insufficient basis on which to base an inference. Students should recognize not only the key role that unusualness plays in the inferential reasoning process but they must also be able to quantify how unusual an outcome is under an initial hypothesis.

Third, the goal of instruction in this study was organized around the idea of tasks as *didactic objects* (Thompson, 2002). A didactic object refers to a thing to talk about that is designed with the intention of creating and supporting reflective mathematical discourse, or in this case, statistical discourse. Thompson (2002) describes how instructional activities are not didactic objects in and of themselves; rather, they become didactic objects when opportunities for conversation and reflection on conceptual issues arise through the goal-directed mental and physical actions of the learner as he participates in a particular instructional task. When instructional activities produce environments that foster reflective goal-directed interactive discourse they become didactic objects; that is, they become opportunities for generating observable information about student understanding. In the context of this study, my goal was to implement the instructional activities as didactic objects; that is, as tools which could be used to produce desirable engagements and conversations between myself and the student.

In addition to the above perspectives there was also a theoretical framing that shaped the design of instruction in terms of the process of inferential reasoning. Following Saldanha (2004), I viewed the structure of the inferential process in terms of an observed sample, an initial hypothesis about the population the observed sample came from, a random procedure for selecting objects from this population, a resulting sampling distribution and an inference from the observed samples location in the distribution back to the population. This conception entails the student conceiving of a sampling event as but a single instance of an underlying repeatable process which over the long run produces a distribution of outcomes which vary naturally around the population parameter. In this view, the empirically produced sampling distribution becomes the basis for quantifying the unusualness of any particular outcome.

Methods and Subjects

In conducting this study, I employed a one-on-one teaching experiment. The teaching experiment methodology as described by Steffe and Thompson (Steffe & Thompson, (2000)) offers the researcher a unique opportunity to bring forth a student's conceptual understandings of particular ideas and the mental operations the student uses to understand those ideas. Moreover, it allows the researcher to experience mistakes in student reasoning. Mistakes that may persist despite the researchers best effort to eliminate them. This allows the researcher to generate and test hypothesis about the boundaries of the students' ways and means of operating. A prime objective within this paradigm, is to document and characterize the subtle shifts in thinking that occur in the context of solving tasks as students' progress from one activity to another and as such it became my methodology of choice.

The participant in this study was Joe[1]--a statistically naïve freshman who had had little or no formal experience with making statistical-based arguments. Joe was recruited from an undergraduate pre-calculus class after an appeal was made to his class for volunteers. Joe participated in 10 sessions in an out-of class setting. Each lesson unfolded over a 75-90 minute period. A written pre-assessment queried his initial intuitions and understandings and a post-activity interview queried his thinking about the key ideas and interconnections among them that were addressed in the designed instructional activities. The teaching experiment itself unfolded

[1] This is a pseudonym

in a sequence of 8 lessons over a two and one-half week period. During the teaching experiment Joe was prompted to explain his thinking both verbally and in written responses in order to gain insight into his reasoning processes. The data corpus includes video-taped discussions around the activity sequence. An analysis of the video produced annotated transcriptions identifying critical events in Joe's reasoning. Joe's utterances were triangulated with his written responses in an attempt to determine the mental actions and ways of thinking that contributed to his understanding things in the way he did. I employed an iterative process of generating and modifying hypotheses in light of the data. All the descriptions and analysis of the Joe's understandings were grounded in his participation in instruction.

Aspects of Instruction

I framed the goal of a statistical inference as being a decision or probability statement about a population that is inferred from a sample. The strategy of such an inference is to consider the long term behavior of samples taken from a given population in order to see if the observed sample is a likely outcome of sampling from that population. This reasoning is a classic example of Fisherian inductive reasoning (Hubbard & Bayarri, 2003) where students assess the strength of evidence against a hypothesis by quantifying how unlikely the observed result would be if in fact the hypothesis was true. This reasoning is similar to a modus tollens-like argument but with an aspect of uncertainty thrown in (Rossman, 2008).

The initial lessons had Joe engage in hands-on resampling activities. These activities involved him drawing an inference from a randomly observed sample to a population. In the later activities, Joe was presented with contextual scenarios that involved the testing of a hypothesis. His task was to physically model the problem; to investigate his models behavior in terms of the samples it produces, and then to interpret the results. This hands-on modeling was then followed by Joe translating the underlying logic of his physical simulation into *ResamplingStats* language using a computer program called *Statistics101*[2]. Joe would use the *ResamplingStats* language to describe the underlying process of his hands-on physical model and then he would run the simulation using *Statistics101* to arrive at an answer to the statistical question. These latter activities shared the following structure: (a) Think through the assumptions of the investigation; (b) Create a simulation model of the situation; (c) Implement the model; (d) Draw a conclusion on the basis of the computer simulated results.

Results

This section focuses mainly on events that transpired over the first five of eight instructional activities. I highlight developments in Joe's thinking that led to the emergence of a robust and non-normative decision rule I've termed the *similarity heuristic*. I also discuss aspects of the logic of hypothesis testing that were especially problematic for Joe; aspects that have to do with the logical connections between the observed sample, the null-hypothesis, and the empirically produced sampling distribution.

Sampling Distribution as a Comparison Device

Activity 1

The instructional intent of Activity 1 was to introduce Joe to the logic of hypothesis testing via coin flipping. I showed Joe what he presumed to be a fair coin, but was actually double-headed. We began flipping the coin and looking at the results. As the heads accumulated Joe

[2] *ResamplingStats* is a simple to use language requiring no previous experience with programming. Each command executes an operation that mimics the random selection of coins, cards or other items.. *Statistics101* (Grosberg, 2012) *i*s a computer program that understands and executes programs written in the *ResamplingStats* language.

quickly realized that something was amiss. After 6 heads in a row Joe indicated that from his experience what was happening was very unusual. After the coin had produced 10 heads in a row Joe wanted to see the coin so he could check it out for himself. Joe perceived that10 heads in a row was possible but exceedingly unlikely if the coin was fair like I told him and so he was highly suspicious that something else besides chance was going on. In this familiar case of coin flipping Joe instinctively understood that we're not interested in what's possible, we're interested in what is likely and hence what the best explanation for the unusual results are. So Joe concluded that the best explanation was that the coin was not fair like I told him. His intuitive reasoning embodied the core logic behind testing a hypothesis and I sum it up in the following argument:

If the coin is fair like the instructor told me, then getting 10 heads in a row is, in my experience, extremely unlikely.

The coin the instructor flipped produced 10 heads in a row.

This is so implausible he must be lying about the coin and so I'm going to reject the notion that the coin is fair (although I'm not absolutely certain).

Activity 2

Activity 2 centered on looking at distributions of sampling outcomes for many samples drawn from a single population in order to judge whether an observed sample is likely to have come from the population that produced the distribution. The general aim was to have Joe view resampling from a known population as a natural strategy for drawing a conclusion about whether an observed sample might have come from that population. By examining the place of the observed sample in the distribution he could judge the degree to which it is a surprising or not surprising outcome.

The scenario was that I placed an opaque bowl filled with 200 candies in front of him and had him randomly select 10 candies. Unbeknownst to Joe the bowl contained 20% red candies and 80% green candies. As it turned out his sample contained 2 reds. The goal was to coax Joe into seeing the logic of making a probability based decision as to whether or not the candy bowl he just sampled from contains 70% red candy. To this end, I showed him how a 70% red candy bowl actually behaved under repeated sampling with respect to the number of reds in samples of size 10 so that he would see how unusual it is for such a bowl to produce the sample he got. The graph below displays the results of 50 random samples of size 10 taken from a candy bowl whose composition was 70% red candy.

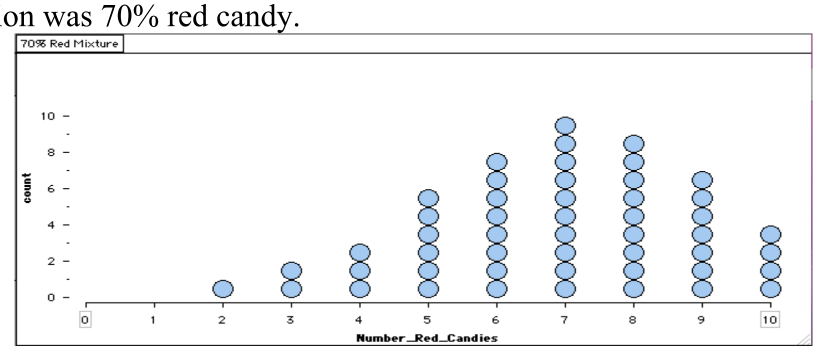


Figure 1. 50 samples taken from a 70% red mixture

Based on this graph I had Joe respond in writing to a series of questions meant to support the notion that, although a random sample is random, what is likely or unlikely to occur in the long run has predictability. The questions I had him answer were meant to foreground the idea that an inference from any individual sample to a population is possible based on the relative frequency patterns that emerge in collections of samples over the long run. In Question 6, Joe was asked to look at the above graph and then think hard about how these simulated results could enable him to decide whether the candy bowl from which he just randomly drew 2 reds in a sample of 10 was likely to have been a 70% red mixture. The question was meant to advance the notion that a sample with just 2 reds is a highly unlikely occurrence from a 70% red candy bowl and therefore a bowl that produces it is most likely not 70% red. The question *did, in fact, not* advance this notion but rather had an altogether different effect than intended and I highlight it because it was the first instance in which Joe displayed a tendency to compare distributions and look for similarity.

In answering the question, Joe's intuitive strategy was to obtain 50 random samples from the opaque candy bowl which produced the sample of 2 reds, calculate the number of reds in each sample, and then compare the resulting distribution to the distribution of the graph I provided. If the candy bowl was indeed 70% red then it should produce a similarly shaped distribution. If the resulting distribution was not sufficiently similar to the one in the graph then he would conclude that the bowl which produced the 2 reds was not 70% red. My immediate concern was that the activity wasn't producing any images in Joe's mind of drawing an inference to the population based on how unlikely it was to see his observed result. As this is the imagery behind hypothesis testing I attempted to redirect his thinking. I asked Joe to imagine drawing 10 candies, two of which are red, from a bowl that I claim contains 70% red candies. This time, however, I pointedly emphasized that he doesn't get to take any more samples from the bowl. He was to make a decision about the claim based solely on his one sample and the graph I provided. After a long silence, Joe replied that the one random sample alone isn't helpful, it's not enough information. He'd have to take more random samples from the bowl that produced the sample in order to see how it distributes itself so he could compare the distributions. Upon further discussion, he remained unpersuaded that the sampling distribution I provided was in any way helpful other than for being a comparison device.

I was hoping Joe would see the relationship between a candy bowl's composition and the random samples it produces and notice that this relationship can be looked at in terms of likely and unlikely samples which cast serious suspicion on the purported composition of another candy bowl when it randomly produces one of the unlikely samples. Hence, he would be positioned to reasonably conclude that the bowl which produced his sample with just two reds was most likely not 70% red. No additional sampling is necessary. These ideas, however, were not forthcoming. While his comparison procedure was not unreasonable, my problem with it was that the observed random sample played no role in determining the relative probability of the initial claim about the candy bowl's composition. Instead of the observed sample being the informational link between the initial claim and what the distribution tells us about the validity of the claim it becomes of no singular importance. I had anticipated that the reasoning which justified his rejecting that the coin was fair in Activity 1, i.e. one highly unusual random sample, would carry over in some respect to Activity.2. At this point I attributed the non-linking of this logic to poor design. I had no reason to think that his comparison strategy would continue into the next activities and eventually morph into a non-normative and robust way of thinking.

The Emergence of a Similarity Heuristic

As a prelude to Activity 3 I had the student once again grapple with how to test the hypothesis that the composition of a bowl of candy was 70% red. His comparison strategy from the previous lesson did not include imagery that supports a logical understanding of hypothesis testing. I was encouraged that under heavy scaffolding Joe seemed to come around and accept the idea that an inference can be made based solely on how likely or unlikely the observed sample was to have come from a 70% red mixture.

Activity 3

In the first part of Activity 3 Joe and I created our first simulation model to test the claim that a candy bowl which produced 2 red candies in a random sample of size 10 was likely to be a 50% red mixture. We began by assuming that the candy bowl was indeed 50% red. We then modeled taking random samples of size 10 from the candy bowl by flipping a coin 10 times. The logic being that since we have assumed the population proportion of red candies in the bowl is exactly half we can let getting a "head" stand for drawing a red candy. Ten flips of the coin then represent drawing one random sample of 10. The number of heads in the ten flips represents the number of red candies in the sample. Then by keeping track of the number of heads in each sample we can estimate how likely it is in the long run to get 2 or fewer red candies in a sample of 10. After explicating this logic, I had Joe give written responses to 6 questions that queried his understanding of the process. After discussing his responses he seemed to understand the logic of what we were doing and so we actually computed a small X-plot of 15 sampling outcomes. I asked him what he would look for if we had the time and patience to continue the X-plot for say 50 samples. He said would observe where "the trend is" and see how often 2 reds popped up. He said that the main trend, or peak of the distribution, should be around 4 or 5, probably closer to 5. I interpreted this explanation as his image of what he was expecting to observe in the empirical data under the initial hypothesis. He also made the rather puzzling comment that if the bowl was coming up with 2 reds a lot of the time then that would be evidence against the claim. As will be seen, the statement was a portent of things to come, but at the time, I simply let the comment pass unexplored. In the second part of the Activity 3 we used a similar reasoning process to investigate the claim that an observed sample with just 2 red candies came from a 40% red candy bowl. This time, however, the random generating device was 10 poker chips, 4 of which he marked red.

Activity 4

In Activity 4 the extent of Joe's conceptual misunderstandings of the logic of hypothesis testing began to surface. Activity 4 was marked by a move to engage Joe in designing a computer simulation using the *ResamplingStats* language in order to test the hypothesis that a random sample of 10 candies with just 2 reds came from a 50% red mixture. The commands in *ResamplingStats* mimic what the hand does when working with a random generating device such as a coin and so Joe and I translated the coin flipping procedure he used in Activity 3 into computer code that would simulate drawing 100,000 random samples of size 10 from a candy bowl whose composition is 50% red while keeping track of the number of red candies obtained each time. The output would be a histogram of the sampling distribution along with the probability of obtaining 2 or fewer reds. Since the chance of obtaining 2 or fewer reds has to be very unlikely if we are to reject the hypothesis, I decided to ask him before running the program for a cut-off criterion for the observed sample below which he would reject the hypothesis. The following exchange then took place[3]

Episode 1, Activity 4

Segment 1

I: You can reject or not reject the hypothesis depending on how unusual your observed sample is. For instance, if the results of running the program indicate that the probability

[3] "I" denoted the instructor's utterances and "J" denotes Joe's utterances

of getting two or fewer reds is less than some cut-off then you can reject the notion that the bowl of candy was 50% red. Does that make sense?

J: Yeah, but actually we can already say it isn't. We're testing for 50% red. That's what the samples are for. We're seeing how many reds out of 10 pop up and how often.

I: Huh? But you're sampling from a bowl you know has 50% red candy right?

J: I am?

I: Yeah, you set it up that way, remember? Now you want to see if it really is 50% red.

J: Oh yeah, I guess I twisted it up.

This excerpt shows that Joe had immediately lost sight of what population he was sampling from and why he was even sampling from it in the first place. The initial assumption that the bowl is 50% red was our working assumption. We set it up for the very purpose that it could be made susceptible to a probability estimate. I *thought* he understood this. This should have set of all types of alarm bells but at the time of this exchange I simply corrected him and moved ahead. I continue.

Episode 1, Activity 4

Segment 2

I: So where do you want to make your cut-off?

J: I have to see the probability first.

I: What? No you don't. You want to state it in advance. What will convince you? I mean, if 2 or fewer reds turn up just 10% of the time, is that strong enough evidence for you?

J: No.

I: How about 5% of the time?

J: Yeah, that'll work.

I: So you want to go with 5% of the time?

J: Oh, I thought you were going in the opposite way. 10% will probably rule it out.

I: So you're going with 10%?

J: Yeah, especially in so many trials.

I: So if the results show that you can expect to get 2 or fewer reds 10% of the time then this will be enough to rule out that the bowl was 50% red?

J: Yeah. No wait. It will be 50% red.

I: What?

J: Yeah, anything less than 10% and I'll say the bowl is 50% red

I: You'll say that it is? Okay wait. I'm not following you.

J: Well from 10% or under it's reasonable that the bowl is 50% red, but if it goes above 10% then I'm going to think it's not 50% red.

I: Say that again.

J: It's like the less reds that show up the more likely the bowl is 50% red because it should be more like 4 to 5 reds being pulled not 2 or fewer. So from 10% down it's reasonable that there would be that many reds showing up and it still be 50-50, but once you start getting 11, 12, 13, 14, 15 percent then you are going to think there are less reds in the bowl and it's not 50% red.

I: I don't get what you're saying.

J: I'm saying that once you start getting 2 reds above 10% of the time then I would reject it because 10% or lower is still reasonable.

I: What if you are getting 2 reds 20% of the time?

J: That wouldn't be reasonable for it to be 50-50.

I: But 20% of the time means that samples with 2 or fewer reds are being pulled every fifth time

J: Yeah

I: So then it's not unusual.

J: It is for a bowl that is 50% red. It's asking us about a bowl that is 50% red and what we would think about it based on these probabilities.

Whatever his reasoning was at this point, it was completely inconsistent and incompatible with the reasoning *I thought* I had been promoting. I felt like we were speaking different languages. For me, a small probability of seeing the observed sample was evidence that the initial hypothesis was not plausible. Joe seemed to be using the probability of seeing the observed sample to make a decision about something being unusual or implausible but it was certainly not the unusualness or implausibility I was talking about. As this took place I was desperately trying to figure out why a large probability of seeing the observed sample would make him reject the idea that the bowl was 50% red? A large probability would mean that under the hypothesis we started with it's actually a common occurrence. I continue with the excerpt.

Episode 1, Activity 4

Segment 3

I: I'm thinking about what you are saying but I'm not getting it.

J: It's like what we did in the other activity with the X-plot. If 10% of the time we're getting 2 or fewer reds being pulled that's like having a couple of X's over the 2. That's not a lot at all so it's reasonable. Most of the results were on the other side and so it supported the fact that the bowl was 50-50. But if there are 15 or 20 percent of the X's over the 2 or smaller area then that looks unusual why that many samples with 2 or fewer were being pulled. That makes me think that the bowl was more like 30% red. So you can reject that the bowl is 50% red if you are pulling 2 or fewer reds a lot of the time.

I: But pulling 2 or fewer reds a lot of the time would make getting a sample with 2 or fewer reds a common occurrence and so you wouldn't reject the hypothesis.

J: No, you would reject it.

I: Why would you reject the hypothesis if the observed sample is a common occurrence?

J: I just told you.

I: But we don't reject the hypothesis if 2 or fewer reds is a common occurrence.

J: This isn't making sense.

I: I guess this is a hard concept.

J: I get it. You're just not getting what I'm saying. Can we please just run the simulation?

I: Explain it to me one more time.

Eventually, after a very long and heated discussion, I began to piece together what Joe was doing. When I asked him to make a cut-off criterion as to what would be a sufficiently unusual outcome he apparently adapted his comparison strategy from the previous lessons to this situation. Recall that Joe's preferred strategy for testing a hypothesis about the composition of a candy bowl in Activity 2 was to repeatedly sample from the bowl that produced the sample and then compare its sampling distribution to the known distribution. If the two candy bowls had sufficiently similar distributions then the claim was accepted, if not, then the claim was rejected. This comparison strategy had now morphed. His decision procedure was now based on comparing the empirically produced sampling distribution to his image of how a 50% red population should be distributing its sampling outcomes. He was literally deciding at what point the empirically obtained distribution would be inconsistent with what he imagined should be produced. In his image of a 50% red population it would be highly unusual if 2 or fewer reds were occurring too often and so that's why he was insisting that a high probability of seeing samples with 2 or fewer reds would be grounds for rejecting the 50% red hypothesis. Never mind that the simulation was set up to sample from a 50% red population so that no matter what the results turn out to be they are results that come from a 50% red population.

Moreover, since all the probability computations for the data assume a 50% red population he was essentially disabled from understanding what the probability output was a probability of. As a result, the key statistical question as to whether obtaining 2 or fewer reds was probabilistically consistent with coming from a 50% red population was not a part of his thought processes. It's no wonder we couldn't understand each other. When we finally ran the simulation, 2 or fewer reds occurred 5.4% of the time which seemed to comport just fine with his subjective image of how a 50% red population should be behaving and so he accepted the hypothesis that the candy bowl was 50% red. In this case his decision was the same that sound reasoning would have led to but the reasoning that got him to this conclusion was completely incompatible with the modus tollens-like logic that supports hypothesis testing. Before ending the lesson that day we had a long discussion in which he seemingly came around to recognizing that the key issue is whether the observed sample is unusual or not under our initial hypothesis. I actually finished Activity 4 with the hope that his similarity heuristic would be a fleeting instance of confusion.

The Similarity Heuristic is a Conceptually Robust Way of Thinking

In Activity 5 through Activity 8 I engaged Joe in conceptualizing textual scenarios as probabilistic situations the goal of which was to make a statistical inference. The activity of designing simulations for the scenarios was intended to force him to come to grips with the underlying logic of the hypothesis testing process. Each activity involved Joe reconstruing the given situation in terms of an assumed population, an observed sample, and a random repeatable sampling process which would allow him to quantify the unusualness of the observed sample under an initial assumption about the population. A primary goal of Activity 5 was to formally introduce Joe to the logical device of a null hypothesis.

Activity 5

In Activity 5, Joe investigated whether right-handed people tend to also be right-eye dominant. According to the scenario he, as the researcher, randomly selected 16 right-handed people and found that 12 of them were right-eye dominant. His null hypothesis was that among right-handers, the right and left eyes are equally likely to be dominant; that is, among right-handed people, 50% will be right-eye dominant and 50% will be left-eye dominant. Joe designed a hands-on simulation using a coin as his random generating device and then we translated his physical simulation into *ResamplingStats* language. Before running the simulation I asked him what he wanted to use as his cut-off criterion. What follows illustrates that his similarity heuristic from the previous activities was a pervasive and conceptually robust way of thinking.

Episode 1, Activity 5,
Segment 1

I: Okay, so suppose out of 1000 trials you get samples with 12 or more say 50 times. That means your observed sample occurred just 5% of the time. Would that be enough evidence for you to reject the null hypothesis?

J: 25% will reject it.

I: What? But that would make your observed sample a common place occurrence.

J: Oh yeah, the other way around, 75%.

I: What?
J: 75%.
I: But if your observed sample is occurring 75% of the time…
J: That would reject the null hypothesis.
I: What?
J: That would reject the null hypothesis because they wouldn't be the same.
I: What wouldn't be the same?
J: The distributions wouldn't be the same.
I: You keep saying stuff like that, but we're taking our samples from a population that we know splits itself 50-50 on the issue, right?
J: But it says 75%. It says 12 or more out of 16 which is 75%.
I: That was the observed sample. We want to see how unusual that is.
J: It's supposed to be 5% right? I remember you saying something about 5%
I: It doesn't have to be 5%. I'm just saying that 5% is typically used. 5% makes the observed sample fairly unusual.
J: But if it happens 5% of the time you wouldn't reject your null hypothesis.
I: Why not? That's where you could reject it. That's why it's typically used.
J: With 5% of the time? But we're saying that it's 50-50, that it's even, and if it's occurring 5% of time it could still be even.
I: It could still be even?
J: Yeah. That's not enough evidence to say it's not 50-50. 5% is like 1 out of 16. But with our data we can reject it because we got 12 out of 16.
I: But 12 out of 16 is the observed sample. We want to see how unusual 12 out of 16 are.
J: So we generate a bunch of samples.
I: Exactly. We generate a bunch of samples from the null hypothesis population and if getting 12 or more right-handers who are right-eye dominant occurs quite frequently…
J: Then you reject the null hypothesis
I: No. Then you don't reject the null hypothesis. If it rarely occurs then you reject the null hypothesis.
J: This isn't making sense.

After further discussion it was clear that Joe was still coordinating his mental image of expected outcomes based on an imagined null hypothesis distribution with empirical data that would either confirm or conflict with that image**.** His decision about the null hypothesis depended on whether he believed the probability of seeing the observed sample or something more extreme in the data was different enough from his expectations to warrant rejecting the hypothesis. The results of repeated sampling were being used to confirm or refute the null hypothesis based on how much they resembled his image of how the outcomes should be dispersing themselves. He was profoundly reluctant to internalize the fundamental logic that the empirical sampling distribution *is* the null-hypothesis distribution—for the very reason that we set it up to be that way. We're not testing to see if it is the null hypothesis distribution, we already know it is. The assumption that the null hypothesis is true is what provides the necessary bridge between the observed sample and a decision about where it came from and without this fundamental understanding nothing I said was making sense to him.

In the remainder of the study, even though his written and verbal responses did at times reflect an understanding of the logic behind employing a null hypothesis, his understanding was fragile and he continued to flit back and forth, more often than not, using his similarity heuristic to decide whether the empirically obtained data conformed to his expectations based on an imagined null distribution. Evidence from the final interview strongly indicated that he had not internalized the scheme of ideas the activities had promoted as the inferential line of reasoning. Because his similarity heuristic seemed to defy remediation, I presented him with the following question during that final exit-interview. I fully expected him to answer it exactly as he did. The question and his response are provided below. They reveal how deeply embedded in his thinking this similarity heuristic was.

> When you run a *Statistics101* simulation program the probability of seeing the observed sample that is output quantifies the strength of evidence against the null hypothesis. Select from the choices below the answer that best fits the inferential reasoning process.
>
> (a) The smaller the probability, the stronger the evidence against the null hypothesis
>
> (b) The larger the probability, the stronger the evidence against the null hypothesis
>
> Explain the choice of your answer
>
> *You are testing how unusual something is and so a larger probability makes it too usual and that's stronger evidence to reject it's the null hypothesis.*

Concluding Remarks

This study engaged a student in designing simulations of sampling experiments as a vehicle for making distributional judgments as to whether an observed outcome was unusual under an initial hypothesis. Physical hands-on sampling procedures were initially used with the intention that they would ground the student's thinking about the kinds of samples that are produced under some initial assumption about a population. After sufficient experience and discussion about this process I conjectured that it would be advantageous to move the student into an environment where the process becomes automated yet the logic remains visible to the learner. I also conjectured that an understanding of the logic behind a hypothesis test would be a necessary condition for the student to construct the necessary commands in order to successfully implement a *Statistics101* program. It's now conceivable that a learner may be able to use both hands-on resampling and simulation using *Statistics 101* to test a hypothesis without fully understanding or being able to coordinate the logic behind the process. Indeed, the student in this study experienced overwhelming difficulty generating and composing the necessary images to support a coherent understanding of hypothesis testing. Moreover, this report characterized how early in the experiment the student structured hypothesis testing around a decision rule based on a judgment of whether the repeated sampling process was or was not producing a distribution sufficiently similar to his imagined null hypothesis distribution. This heuristic became so deeply ingrained in his thinking that it conceptually disabled him from operationalzing the logic of hypothesis testing. While it is admitted that a sample of size one does not support making claims about the prevalence of this study's findings it is, nevertheless, an unbefore documented existence proof of how a student might reason.

The student's unsound similarity heuristic when taken together with previous research suggests that hypothesis testing may be an inherently difficult set of concepts that invite misunderstanding regardless of how they are taught. A large part of the difficulty seems to be rooted in the indirect modus tollens-like reasoning which connects the null hypothesis to the probability estimates that follows from it. People do not set up null hypotheses to test assertions in their everyday reasoning and so the statistical form of this logic is very unnatural.

During this study I encountered at least 5 fundamental notions that a student's attention must be kept focused on if he is to connect the ideas that hold the logic of hypothesis testing together. If any of the following statements seem obscure, unclear, or confusing then the student will probably not assimilate the logic behind the process and may fall victim to a fallacious form of reasoning as the student in this study did.

1. The null hypothesis is a working assumption about where the observed sample came from.
2. The hypothesis test begins by assuming the null hypothesis is true.
3. The data collected by resampling tells us what types of samples are likely or unlikely under the null hypothesis
4. If the observed sample is very unlikely under the null hypothesis, then we can reasonably conclude that the null hypothesis is not an accurate description of the population from which the observed sample came.
5. If the observed sample is not unexpected under the null hypothesis, then one must conclude that there is insufficient evidence for rejecting the null hypothesis as an accurate description of the population from which the observed sample came.

The results indicate that the instructional activities generally failed to help the student develop his reasoning about hypothesis testing in the way I intended he would. This highlights the challenges of creating a sequence of activities that will perturb a student's reasoning as intended and result in the student developing a targeted understanding.

Reconsidering the design of the Activities

In retrospect, I speculate that much of the Joe's confusion could have been avoided or at least kept to a minimum by suppressing the technical term "null-hypothesis" until the very end of the study. He almost always misstated it and he never seemed to grasp why we were even stating it in the first place. This led to confusion about what the empirical sampling distribution was even a distribution *of* and what the probability of the observed data was supposed to be used for. For intuitiveness and ease of exposition I conjecture that it might be helpful to initially introduce hypothesis testing as a debate between a Skeptic and an Advocate (Wardrop, 1995). The Skeptic always sees any observed outcome as due to chance alone. The Advocate, on the other hand, acknowledges that the Skeptic could possibly be correct but argues that the Skeptic's conclusion strains credibility and that the observed outcome is unlikely to be due to chance. This leads to a hypothesis test and the question becomes, who will win the debate? After collecting data via simulation--always remembering that the sampling distribution is generated on the assumption that the Skeptic is correct—the student obtains a number called the P-value. The P-value is presented as measuring how likely it would be to see the observed sample or something more extreme if the Skeptic is correct. The smaller the P-value the more likely the Skeptic is wrong. I conjecture that this approach merits future study as an entry way into the logic of hypothesis testing and would likely best be done with a team of students who could role play as the Skeptic and the Advocate.

References

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.

Chance, B. L., & Rossman, A. J. (2006). *Using simulation to teach and learn statistics.* Paper presented at the Proceedings of the seventh international conference on the teaching and learing of statistics, Salvador, Bahia, Brazil.

Cobb, G. (2007). The Introductory Statistics Course: A Ptolemaic Curriculum? *Technology Innovations in Statistics Education, 1*(1).

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).

Garfield, J., & Ben-Zevi, D. (2008). Learning to reason about statistical inference. In J. Garfield & D. Ben-Zevi (Eds.), *Developing Students' Statistical Reasoning*: Springer.

Grosberg, J. (2012). Statistics101: Statistics the easy way! Resampling, Bootstrapping, etc., from www.statistics101.net/

Hubbard, R., & Bayarri, M. J. (2003). Confusion over measures of evidence (p's) versus errors (a's) in classical statistical testing. *The American Statistician, 57*(3).

Kaplan, D. (2007). Computing and Introductory Statistics. *Technology Innovations in Statistics Education, 1*(1, Article 5).

Lipson, K. (2002). *The role of computer based technology in developing understanding of the concept of sampling distribution*. Paper presented at the Proceedings of the sixth international conference on the teaching of statistics. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/1/6c1_lips.pdf

Mills, J. D. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of statistics education, 10*(1).

Rossman, A. J. (2008). Reasoning about informal statistical inference: One statisticians view. *Statistical education research journal, 7*(2).

Saldanha, L. A. (2004). "Is this sample unusual?": An investigation of students exploring connections between sampling distributions and statistical inference. Unpublished Ph.D Dissertation. Vanderbilt.

Saldanha, L. A. (2011). *Aspects of conceiving Stochastic experiments.* Paper presented at the The 35th conference of the international group for the pyschology of mathematics education

Ancara, Turkey.

Saldanha, L. A., & Thopmson, P. W. (2007). Exploring conncections between sampling distributions and statistical inference: an Analysis of Student's engagement and thinking in the context of instruction involving repeated sampling. *International electronic jouranl of mathematics education, 2*(3).

Steffe, L., & Thompson, P. (2000). Teaching Experiment Methodology: Underlying Principles and Essential Elements. In A. E. Kelly & R. Lesh (Eds.), *Research design in mathematics and science education* (pp. 267-307). Hillsdale, NJ: Erlbaum.

Thompson, P. (2000). Radical Constructivism: Reflections and Directions. In L.P.Steffe & P. W. Thompson (Eds.), *Radical Constuctivism in action:Building on the pioneering work of Ernst von Glasersfeld* (pp. 412-448). London: Falmer Press.

Thompson, P. W. (1994). The develpment of the concept of speed and its relationship to concepts of rate *The develpment of multiplicative reasoning in the learning of mathematics* (pp. 179-234). Albany: SUNY Press.

Thompson, P. W. (2002). Didactic objects and didactic models in radical constructivsm. In K. Gravemeijer, R. Lehrer, B. V. Oers & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 191-212).

Von Glasersfeld, E. (1995). *Radical Constrictivism: A way of knowing and learning*. London Falmer Press.

Wardrop, R. (1995). *Statistics:Learning in the Presence of Variation*. Dubuque, Iowa; Melbourne, Austrailia; Oxford, England: Wm. C. Brown.

Yu, C. H. (2007). Resampling: A Conceptual and Procedural Introduction. In J. Osborne (Ed.), *Best Practices in Quantitative Methods*: Sage Publications.

FACTORS INFLUENCING STUDENTS' PROPENSITY FOR SEMANTIC AND SYNTACTIC REASONING IN PROOF WRITING: A SINGLE-CASE STUDY

Juan Pablo Mejia-Ramos[1] Keith Weber[1] Evan Fuller[2] Aron Samkoff[1] Kevin Calkins[1]
[1] Rutgers University [2] Montclair State University

In this paper we present a case study of an individual student who consistently used semantic reasoning to write proofs in calculus but infrequently used semantic reasoning to write proofs in linear algebra. We argue that the differences in these reasoning styles can be partially attributed to this student's familiarity with the content, the teaching styles of the professors who taught him, and the time he was given to complete the tasks. These results suggest that there are factors that have been ignored in previous research, including domain, instruction, and methodological constraints, that researchers should consider when ascribing to students a proving style.

Key words: Proof; Proving styles; Semantic proof productions; Syntactic proof productions.

1. Introduction

In recent years, several mathematics educators have noted that one can successfully engage in advanced mathematics in two different ways. An individual can focus on the formal aspects of mathematics by understanding statements in terms of their logical structure and the formal definitions of their terms, and by using calculation and formal rules of inference to produce new mathematical ideas. Alternatively, an individual can understand and reason about mathematical concepts using informal representations of these concepts, such as thinking about these concepts in terms of graphs, diagrams, or examples (Pinto & Tall, 1999, 2001; Raman, 2003; Vinner, 1991; Weber & Alcock, 2004, 2009). Recently a number of research reports have identified individual students or groups of students who predominantly engage in one of these two types of reasoning while rarely engaging in the other (e.g., Alcock & Inglis, 2008; Alcock & Simpson, 2004, 2005; Alcock & Weber, 2010; Duffin & Simpson, 2006; Moutsioz-Rentzos, 2009; Pinto & Tall, 1999, 2002; Weber, 2009). However, in these studies, these students' reasoning styles were often identified by their performance in a small number of tasks and these tasks were nearly always situated in a single mathematical domain. In this paper, we present a case study of an individual with very different ways of reasoning about linear algebra and calculus tasks. We use this case study to illustrate that an individual's reasoning in advanced mathematics is not necessarily consistent across domains, and we suggest factors that should be taken into account when considering a student's reasoning style.

1.1. Different modes of reasoning in advanced mathematics

The products of advanced mathematical reasoning (definitions, conjectures, theorems, and proofs) are typically expressed in a unique representation system. This representation system uses a combination of specialized words and logical syntax, where the nouns and adjectives have precise definitions that are agreed upon by the wider mathematical community, and the rules of inference are based on logical deduction (Weber & Alcock, 2009).

One way that mathematics can be understood is working within this system. For instance, a concept can be understood by studying its definition, making deductions from this definition to see what properties are true about that concept, reformulating the definition into equivalent statements, and comparing the logical structure of the definition to other concepts that one is aware of (e.g., Pinto & Tall, 1999; Weber, 2009). When writing proofs, one can start with definitions and permissible assumptions and use logical inference (including applying known theorems) and calculation to deduce the desired conclusion without considering informal representations of that concept such as graphs or diagrams (e.g., Vinner, 1991; Weber & Alcock, 2004). Weber and Alcock (2009) refer to this type of reasoning as syntactic reasoning.

Alternatively, a concept can be interpreted in other representation systems, by relating the concept to graphs, diagrams, or prototypical examples. An individual may try to understand a concept by constructing links between the definition of the concept and these informal representations, perhaps refining their informal understanding if necessary (e.g., Pinto & Tall, 1999). One can write proofs about these concepts by constructing informal arguments using these representations and using these informal arguments as a basis for constructing a formal proof (e.g., Raman, 2003; Weber & Alcock, 2004). Weber and Alcock (2009) refer to this type of reasoning as semantic reasoning.

1.2. Cognitive styles in advanced mathematics

To define *cognitive styles*, we adopt the perspective of Riding and Cheema (1991), as used in mathematics education by Duffin and Simpson (2006). A *cognitive strategy* is a general cognitive approach that an individual can use to accomplish a class of mathematical tasks. An individual is said to have a *cognitive style* if he or she consistently invokes the same cognitive strategy when working on a class of mathematical tasks. Semantic and syntactic reasoning constitute different cognitive strategies that an individual may use to engage in advanced mathematical thinking. Recently, several researchers have suggested that students and mathematicians may have cognitive styles in advanced mathematics—that is, they may show a strong propensity to consistently engage in syntactic or semantic reasoning.

Pinto and Tall (1999) hypothesized that this may be the case with mathematicians, citing Poincare (1913), who referred to Reimann as an intuitive thinker “who calls geometry to his aid” and Hermite as a logical thinker who “never invoked a sensuous image” when discussing mathematics (Pinto & Tall, 1999, p. 281). Similar contrasts can be found when mathematicians reflect on their own practice. For instance, Andre Weyl (1940) claimed that, when doing mathematics, a mathematician “forgets what the symbols stand for… there are many operations that he can carry out with these symbols, without ever having to look at the things they stand for”. In contrast, William Thurston (1994) stated that when doing or reading about mathematics, he is continually trying to interpret ideas in terms of his mental models that are not represented by formal mathematical syntax. In a large-scale interview study, Burton (2004) identified some mathematicians whose reasoning was predominantly symbolic and others whose reasoning was mostly visual. In mathematics education, researchers have also hypothesized that, like mathematicians, students may have cognitive styles in advanced mathematics (e.g., Alcock & Inglis, 2008; Alcock & Simpson, 2004, 2005; Alcock & Weber, 2010; Duffin & Simpson, 2006; Moutsioz-Rentzos, 2009; Pinto & Tall, 2002; Raman, 2003; Weber, 2009), and have begun to explore what makes students with either reasoning style successful or unsuccessful.

1.3. Difficulties with documenting cognitive styles

In a comprehensive synthesis of the research on cognitive styles, Coffield, Moseley, Hall, and Ecclestone (2004) concluded that both researchers and teachers frequently label individuals as having a particular cognitive style based on insufficient evidence. They warn that one should not infer that students have a cognitive style due to their behavior completing a small number of tasks, or tasks within a narrow setting. The cognitive strategy that one invokes may be dependent upon the demands of the task, as well as other cultural or environmental factors (Entwistle, 1998).

When assigning cognitive styles to students in advanced mathematics, mathematics educators have generally made this judgment based on students’ behavior on tasks within a single mathematical domain. From the research articles that we are aware that assign a cognitive style to students in the context of undergraduate or graduate proof-oriented mathematics courses, we noticed that only Moustioz-Rentzos (2009) analyzed students' cognitive strategies in two different mathematical domains before assigning to them a cognitive style. Further, while Alcock and Simpson (2004, 2005) are careful to indicate that their findings are specific to real analysis, others are less specific about the limitations of their findings and often imply they are discussing students' use of cognitive strategies in advanced mathematics in general.

1.4. Purposes of this paper

In this paper, we present the case study of one student, Caleb, and examine his cognitive strategies for writing proofs in linear algebra and calculus. Caleb consistently applied semantic reasoning while writing proofs in calculus; however, his use of semantic reasoning was much less common in linear algebra, only being exhibited occasionally. In addition to providing an existence proof that students' cognitive styles may be dependent upon the mathematical domain, we use our data to hypothesize causes for the differences in Caleb's behavior and suggest factors that researchers should attend to when assigning reasoning styles to students.

2. Methods

2.1. Materials

This case study comes from a larger study on students' proving processes. In this study, 12 students were asked to write seven proofs in linear algebra and seven proofs in calculus. In each mathematical domain, we included tasks that we judged to be *semantic* if they invited[1] the student to take a semantic approach to prove the statement, *syntactic* if they invited the student to take a syntactic approach to proving the statement, and *neutral* if the task invited the student to take either a semantic or syntactic approach. We also classified tasks as *easy, medium*, or *hard* based on our perception of how many math majors would make significant progress or complete the task. Each task was assigned two labels to describe both difficulty and the type of reasoning it invited. For instance, the "semantic-hard" task was a task with a high level of difficulty that invited a semantic approach. These judgments were based on interviews with mathematicians, the authors' own experience doing and teaching mathematics, and feedback from an advisory board providing guidance on this project. For level of difficulty, we verified our judgment by giving these task items to two classes of undergraduate math majors and examining the proportion of students who were able to complete or make significant progress on the task. The 14 tasks that we used, as well as their labels, are given in the Appendix.

2.2. Procedure

Twelve advanced undergraduate mathematics majors from a research-intensive university agreed to participate in this study and were paid a moderate fee for their participation. Each participant met individually with an interviewer for two 90-minute sessions. Participants were told that: (1) they would be asked to write proofs and to "think aloud" as they constructed the proofs, (2) they would be given ten minutes to complete each proof and, (3) they should write up their final proofs as if they were going to be graded in a mathematics exam.

In the first interview session, the participant began by completing a practice problem to become accustomed to the interview format. The participant was then given one of the study tasks. The participant was permitted to work on a proof until he or she wrote a proof that he or she was satisfied with, the participant felt that he or she could not make any more progress,

[1] Naturally, the invitations were implicit. That is, the task afforded an easily accessible or successful semantic or syntactic approach.

or ten minutes elapsed (whichever possibility occurred first). The interviewer then asked the participant questions about their proving process, including a summary of what the participant did, what the main ideas of the proof were, and how the main ideas of the proof were generated. This process was repeated six more times for other tasks.

In the second interview session, the participant attempted the remaining seven tasks using the same protocol as above. The participant was then asked general questions about his or her proving process, knowledge of linear algebra and calculus, and experience as a student in courses in these areas. All interviews were videotaped.

At any point in the study, two resources were available to the participants. First, if participants could not recall the definition of a relevant concept, they could ask the interviewer for the definition. At that point, the interviewer would hand them a sheet of paper with the definition of the concept and an example of the concept. For the definition sheet of singular matrices, a theorem giving a small set of conditions equivalent to being an invertible matrix was also provided. Second, participants had access to a computer with a graphing calculator that enabled participants to view the graph of any function that they wished.

2.3. The case study of Caleb

In this paper, we report on the performance of one student, Caleb. Caleb was a mathematics major who was studying to be a secondary mathematics teacher. At the time of the study, he had just completed an undergraduate degree in mathematics and was beginning the fifth year of a five-year mathematics education program that leads to a master's degree in mathematics education and state certification to teach secondary mathematics. Caleb was chosen primarily because his approach to different tasks presented an interesting contrast. His proving strategies for the linear algebra tasks and the calculus tasks were very different and, when answering questions about why this might have occurred, he was articulate in describing interesting reasons for these differences. In the first interview session, Caleb completed the seven linear algebra tasks. In the second interview session, he completed the calculus tasks.

2.4. Analysis

We conducted our analyses at three levels of granularity. First, Alcock and Inglis (2009) argued that when evaluating whether syntactic or semantic reasoning was used in a proof construction, what is most important is to determine if the prover represented a concept in a different representation system than the formal mathematical system used in mathematical proof. For this analysis, we flagged every instance in which Caleb drew a diagram or graph, or constructed a specific example of a general concept.

Second, Weber and Mejia-Ramos (2009) argued that it is not only important to know *if* a diagram, graph, or example was considered by a student when attempting to construct a proof. Analyzing *how* these informal representations of concepts were used, including the participants' intended purpose for considering these representations and how insights gained from these representations related to the final proof that was produced can also provide useful information. In this second level of analysis, for each informal representation of a concept that Caleb considered, we describe why he used this representation and how (if at all) it aided his proof construction.

Third, when researchers claim that students have a semantic or syntactic reasoning style in advanced mathematics, they often base these findings as much on interview data as on performance on mathematical tasks (e.g., Alcock & Weber, 2010; Duffin & Simpson, 2006; Pinto & Tall, 2002; Raman, 2003; Weber, 2009). For Caleb, we document patterns of reasoning he used in his calculus and linear algebra tasks and report on interview data when he was asked how he commonly approached these proofs.

3. Results

3.1. Linear algebra tasks

Results of the first two stages of analyses on Caleb's linear algebra tasks appear in Table 1.

Use of informal representation	Contribution to proof attempt
Syntactic-Hard	
Caleb attempted to find a general expression for A^3 when A was a 2x2 matrix.	Caleb announced he was hoping to find a special property A would have if $A^3 = 0$, but never found a general form for A^3 so his investigation yielded no contribution.
Neutral-Hard	
Caleb multiplied two 2x2 matrices to check that det(ST) = det(S)det(T) was correct.	Caleb successfully verified that a theorem that he recalled, det(ST)= det(S)det(T), was correct in these instance. This increased Caleb's confidence in the theorem, although the theorem was not useful for proving the task.

Table 1. Caleb's use of informal representation in Linear Algebra tasks

Use of graphs, diagrams, and examples. As Table 1 illustrates, Caleb did not draw or discuss a graph or a diagram in any of his seven proof attempts. He considered examples in two of the seven proof attempts.

The Syntactic-Hard task asked Caleb to prove that if a square matrix A had the property that $A^3 = 0$, then A had no non-zero eigenvalues. Caleb conjectured that a square matrix whose cube is the zero matrix might have a special property that was useful to this proof. In an attempt to figure out what that special property was, Caleb produced an arbitrary 2x2 matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and began using matrix multiplication to compute A^3, but quickly abandoned this attempt before completing the calculation. He did not verbalize any inferences from his work. We coded this as using an example since he looked at a 2x2 matrix, where the A in the problem statement could have been any square matrix.

The Neutral-Hard task asked Caleb to prove that if a 2x2 matrix T was not invertible, then there existed a 2x2 matrix S such that $TS = 0$. When Caleb began his work, he asked for the definition of invertible. The sheet containing the definition contained the theorem giving several equivalent conditions to a matrix being invertible. Caleb attempted to use the fact that since T was not invertible, det(T) = 0. He then recalled the theorem that det(TS)=det(T)det(S), but was not sure if he recalled the theorem correctly. To verify the theorem, he chose two 2x2 matrices and multiplied them together, writing: $\begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 16 & 17 \\ 6 & 7 \end{pmatrix}$. He then verified that this computation satisfied det(TS) = det(T)det(S), providing him with confidence that this theorem was true. However he was unable to make any perceptible progress on the proof with the use of this theorem.

Qualitative evidence from Caleb's comments and general behavior. Weber and Mejia-Ramos (2009) noted that an absence of semantic reasoning in a student's mathematical work does not necessarily imply a preference for syntactic reasoning, as the student may not have engaged in much reasoning of any type. To claim a student has a syntactic proving style, it is necessary to not only show that semantic reasoning was uncommon, but that the student engaged seriously with the proof production task and exhibited syntactic reasoning. We illustrate how Caleb's work on the linear algebra tasks satisfied both criteria, first by describing his work on two of the tasks and then giving a general account of his behavior and interview responses.

We first illustrate Caleb's reasoning by presenting his work for the Neutral-Easy task, where Caleb was asked to prove that $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x, y, z \in R, xyz \geq 0 \right\}$ was not a subspace of R^3. When Caleb first read the problem statement, he wrote down the definition of subspace, saying that he "starts proof by writing down what [he] knows". He wrote "W is a subspace of R^3 iff for any two vectors $u_1, u_2 \in R^3$, $(\forall c \in R)(\forall d \in R)(cu_1 + du_2 \in R^3)$". He then announced: "to prove it's not a subspace, I need to find one instance where this wouldn't work out [...] what I can do is pick negative scalars and just show that this is not going to be in it". He then deduced that if $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ and $\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$ are in W, then $-\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} -x_1 - x_2 \\ -y_1 - y_2 \\ -z_1 - z_2 \end{pmatrix}$ is not in W. His final proof attempt consisted of verifying that the product of the terms in the final vector was a negative number[2]. Caleb used syntactic reasoning throughout this proof construction—applying the definition of subspace to see what W being a subspace in R^3 would apply in set theoretic language, negating the statement to see how this would be false, and using algebra to verify that he had found a class of counterexamples. Although a single counterexample would have sufficed, Caleb instead used deduction to produce what he believed was a class of counterexamples.

For the Syntactic-Hard task, Caleb was asked to prove that a particular subspace U of R^4 had dimension 2. He began his proof attempt by asking for the definitions of subspace and dimension, and then inferring that any linear combination of vectors in U was still contained in U. He struggled with how he should proceed, saying, "I want to go very mechanical with the definitions but there's so much floating around in my head that I can't get a focal point". He then focused on what he needed to prove, writing "Prove that dim(U) = 2" and saying "this is where I get very mechanical". He reinterpreted the claim to be proven as "dim(U) = 2 means that there are two vectors in any basis of U". He was unsure of how to proceed, and tried representing his previous work as "B = $\{v_1, v_2\}$ is a basis for U". He was unable to make further progress on his proof attempt. When asked to describe his thought processes, he said, "I would almost write that U is a subspace down here [referring to the bottom of the page where he was writing his proof] and try to make them meet". And then, describing why he focused on the statement to be proven, Caleb explained:

> *Caleb:* [...] sometimes, if I do that, and I have this part [referring to the conclusion], that's kind of like the end of the proof on one side, and the beginning of the proof, and I kind of want to make them meet somewhere in the middle.

In general, Caleb worked hard on each of the seven tasks, indicating that his lack of semantic reasoning was not solely due to lack of work of any kind. Across the seven tasks, Caleb spent the full ten minutes working on five of them. For the two remaining tasks (Neutral-Medium and Semantic-Hard), Caleb spent six minutes working on these tasks before abandoning his proof attempt.

Caleb's initial approach to each of the seven linear algebra tasks was to write down the definitions of the concepts involved in the problem, either by recalling these definitions or

---

[2] We note here there is a flaw in Caleb's argument, as x_1, y_1, and z_1, as well as x_2, y_2, and z_2, need not all be positive, so $-x_1$-x_2, $-y_1$-y_2, and $-z_1$-z_2, need not all be negative. If Caleb had specified had arguments only pertained to elements in *W* with three positive terms, his argument would have been valid and produced a general class of counterexamples.

asking the interviewer for the sheet containing the definitions. We also thought it was interesting that at two points during the interview, when Caleb was struggling writing proofs, he complained that what he lacked was experience syntactically manipulating the definitions. At one point, Caleb commented, “I’m not so familiar with the definitions. One of the things that makes proof writing easier is manipulating the definitions again and again”. At another point, he lamented: “the notions are familiar. It’s the routine manipulations that I would have picked up if I had worked with them [the concepts] more”. Hence, there is evidence that Caleb did consistently employ syntactic reasoning for each of these tasks.

Summary of Caleb’s reasoning strategies for linear algebra tasks. Based on his behavior on the seven proving tasks, we would classify Caleb as a student with a propensity for syntactic reasoning who would also rarely engage in semantic reasoning. Even though Caleb worked extensively on seven linear algebra proof construction tasks, he only considered informal representations of mathematical concepts for two of them and these informal representations played only a limited role in Caleb’s reasoning. More often than not, Caleb’s proof attempts predominantly consisted of syntactic reasoning. Most of Caleb’s proof attempts began with him writing the definitions of the concepts involved in the statements to be proven and drawing basic deductions from these statements, at one point noting that “once again I go mechanical to a certain extent”.

3.2. Calculus tasks

The result of the first two stages of analyses on Caleb’s calculus tasks is presented in Table 2.

Use of graphs, diagrams, and examples. As Table 2 illustrates, Caleb considered an informal representation for all seven of the calculus tasks that he attempted. Further, Caleb’s graphs and examples played a significant role in his proof attempts. In four instances (Neutral-Medium, Semantic-Hard, Neutral-Easy, and Neutral-Hard), Caleb graphed functions satisfying the hypotheses of the statement that he was trying to prove, used insights from these graphs to intuitively explain why the statement was true, and then used (or attempted to use) this informal explanation as a basis for his proof. For instance, when completing the Neutral Medium task, immediately after reading the problem statement, which asserted that a function with $f(0) = f'(0) = 1$ and $f''(x) > 0$ for all positive x would have $f(2) > 2$, Caleb immediately began drawing a graph, saying:

> *Caleb:* f(0) = 1 [plots the point (0, 1) on a coordinate system], f’(0) is 1 so the slope right here is going to be a 1 [draws a short line segment around (0, 1) with a slope of 1] and f’’(x) > 0 for all positive x and that means that from here [referring to (0, 1)], the slope is only going to get steeper, so from the y-axis onwards, it’s only going to be concave down [sic], and they’re saying prove that f(2) is going to be greater than 2 … if this was a straight line [draws a straight line with a slope of 1 from (0, 1)], then we’re going to get f(1) is 2 and f(2) is 3 … if this was just a straight line, which it can’t be, then we would have (1, 2) and (2, 3) [plots (1, 2) and (2, 3)]. So, okay, to go by proof, hopefully I can make this rigorous enough.

Our interpretation of this excerpt is claiming that since, if $f(x)$ was a line, $f(2) = 3$, but since *f* was concave up $f(2)$ would have to be greater than 3. Shortly after giving this explanation, Caleb explained, “I can prove it graphically, but I just want to be sure that this would be acceptable to a professor in real analysis”. He proceeded to write a proof by contradiction using the Mean Value Theorem. When asked to give a general description of what he was thinking when working on this proof, Caleb replied, “I was thinking how was I going to translate my graphical understanding into a rigorous symbolic proof that would be accepted by a professor”.

As another illustration, for the Semantic-Hard task, Caleb was asked to show that the equation $x^3 + 5x = 3x^2 + \sin x$ has no non-zero solutions. Caleb recognized this was equivalent to showing $G(x) = x^3 - 3x^2 + 5x - \sin x$ only had a root at zero. Caleb first graphed $f(x) = x^3 - 3x^2 + 5x$ using the graphing application in the provided computer, saw that $f(x)$ only had a root at zero, and then verified this result algebraically. However, he remarked "that sine could do something and give this [the general equation] a second solution theoretically". Caleb then graphed $f(x)$ along with $g(x) = \sin x$ using the graphing application, and noted that "$f'(x)$ is greater than $h'(x)$ because I see it on the graph". He then gave the following intuitive explanation for why the problem statement was true:

> *Caleb:* I graphed these two functions and my only issue was that whatever values the sine function can take on, as long as it can't compensate for how fast the $g(x)$ was growing, then it wasn't going to produce any more solutions. So in my head, I had these two [$f(x)$ and $g(x)$] and that was giving me the shape of the third graph [$G(x)$].

Caleb ran out of time and gave a proof that claimed $f'(x) > g'(x)$ but did not justify this claim. Describing his proof-writing processes, Caleb said, "once again it was a process of me turning my explanatory kind of proof into a symbolic kind of proof."

Use of informal representation	Contribution to proof attempt
Neutral-Medium	
Caleb graphed a linear function that satisfies f(0) = 1 and f(1) = 2.	Caleb observed that a linear function would satisfy f(2) = 3 and deduced that f(2) > 3 for a concave up function. He used this as the basis for the proof.
Semantic-Hard	
Caleb graphed $f(x) = x^3-3x^2+5x$.	Caleb observed that the only solution for f(x) = 0 is x = 0.
Caleb added the graph of h(x)=sin x onto the same set of axes.	Caleb noted that f(x)=h(x)=0 and f'(x) > h'(x) and used this as the basis for his proof
Neutral-Easy	
Caleb drew a prototypical graph of an even function symmetric about the y-axis.	Caleb formulated an argument for why the slopes at x and –x are negatives of one another and used the graph to guide the algebraic work that constituted his proof.
Syntactic-Medium	
Caleb plugged in numbers to explore the claim that $a^2 + b^2 > ab$.	By plugging in numbers, Caleb observed that ab would only be negative if a and b had opposite signs, so he only needed to consider cases where a and b had the same sign.
Neutral-Hard	
Caleb graphed f(x) and g(x) fitting the hypotheses of the task.	Caleb used the graph to see why the statement was true. He later expressed a desire to use the Mean Value Theorem as a tool in his proof and used the graph to see how MVT would be used, but could not finish his proof because he ran out of time.
Semantic-Medium	
Caleb graphed sin x.	Caleb realizes sin x is an odd function and hypothesized that $\sin^3 x$ is odd because the product of odd functions is odd.
Caleb graphed $\sin^2 x$.	Caleb realized the above hypothesis is false, but then set as a sub-goal to prove that the product of two odd functions is an even function and the product of an odd and even function is odd.
Syntactic-Hard	
Caleb unsuccessfully tried to graph a generic f(x) and g(x) = f(x)/x since he could not graph functions satisfying all the hypotheses.	As Caleb could not model the situation, this was not helpful in his proof attempt.

Table 2. Caleb's use of informal representations in Calculus tasks

For two other proofs, Semantic-Medium and Syntactic-Medium, Caleb's exploration of specific example objects provided him with insights that shaped his successful proof production. For instance, for the Semantic-Medium task, Caleb had conjectured that he could deduce that $\sin^3 x$ was an odd function from $\sin x$ being an odd function using the warrant that the product of odd functions was odd. However, upon inspecting the graph of $\sin^2 x$ and realizing that this was an even function, Caleb changed his conjecture to the product of two odd functions is even and the product of an odd function and an even function is odd.

Qualitative evidence from Caleb's general behavior. For five of the seven tasks that Caleb attempted, shortly after reading a description of the problem statement, Caleb attempted to model the statement graphically. If he was successful in doing so, he would try to intuitively explain why the statement to be proven was true and then express this informal explanation using the language and deductive structure required of a mathematical proof. When asked to summarize his proof construction attempts, for five tasks Caleb declared that his proof writing consisted of trying to formalize what he saw in pictures. For instance, he said, "once again I was just thinking how I can bridge the gap between what I clearly saw on the graph into symbols" and "the entire time I saw it on the graph, plain as day, but putting it into symbols is a translation process". Two similar quotations are given in the preceding sub-section. Caleb's proving processes, as well as his descriptions of them, indicate that Caleb attempts to write proofs in calculus using what Raman (2003) referred to as a "key idea", where the activity of proving can be viewed as building informal justifications of mathematical claims and then translating these into the language and structure of formal proof.

Summary of Caleb's reasoning on calculus tasks. We contend that the evidence above shows Caleb as having a strong semantic reasoning style for writing proofs in calculus. For all seven tasks, Caleb considered informal representations of the mathematical concepts and, in each case, these allowed Caleb to draw useful inferences that shaped his proof attempt. Further, Caleb's actions and his reflections upon them indicate he views the task of proving as generating intuitive graphically-based arguments and translating them into the language of proof.

4. Comparing Caleb's behavior on the linear algebra and calculus tasks

We believe Caleb's propensity to engage in semantic reasoning differed depending on whether Caleb was working on linear algebra and calculus tasks. We contend that Caleb had a semantic proving style for the calculus tasks but engaged in semantic reasoning only occasionally for the linear algebra tasks. The differences in Caleb's behavior are summarized in Table 3. This result is theoretically interesting in itself in that it informs mathematics educators that proving style might not solely be a function of the student, but may also depend on the mathematical domain being studied.

In this section, we use Caleb's comments to speculate on why his reasoning strategies were so different. One possibility is simply that calculus is more amenable to semantic reasoning. Just as we would expect almost any students to use diagrams more when writing proofs in Euclidean geometry than in, say, number theory, perhaps most students would use semantic reasoning more in calculus than in linear algebra. Although we cannot rule out this possibility, we use Caleb's response to our interview questions to argue why we do not believe this was the case. Rather, we believe Caleb's behavior should be attributed to other factors, including how he was taught calculus and linear algebra, his depth of knowledge of these two domains, and the time pressure he felt while completing these tasks.

Behavior	Linear Algebra Tasks	Calculus Tasks
Tasks in which an informal representation of an object was constructed.	2	7
Tasks begun by writing definitions of concepts in the statement to be proven.	7	1[3]
Tasks begun by attempting to model the statement to be proven.	0	5
Tasks in which an informal explanation of the statement was provided.	1	4
Description of his proving process.	"When there's a lot of stuff like this, I go mechanical to a certain extent"	"Once again it was a process of me turning my explanatory kind of proof into a formal kind proof"

Table 3. Differences in Caleb's behavior in Linear Algebra and Calculus tasks

4.1. Caleb's proving strategies in other domains

After Caleb completed the calculus tasks, the interviewer noted that Caleb often used graphs to complete the calculus tasks and asked if this was something that Caleb did in general. Caleb responded "yes, I think very graphically" adding that in calculus, he used graphs both to interpret statements and as a basis for his proofs. He was later asked if his more syntactic approaches to writing proofs in linear algebra were because of the differences in the domains of linear algebra and calculus or his lack of familiarity with the linear algebra material (discussed more below). Caleb replied:

> *Caleb:* I really do think that it's because of the lack of familiarity and not the difference in subject area cause if I'm trying to imagine writing a proof about abstract algebra, which I did take the class and I enjoyed it so I got into it a little, I more often than not for an abstract algebra proof am not willing to just put the definitions down but rather to kind of like sink my teeth into it and understand it in my own constructive way and then go about trying to make the proof.

These results suggest that Caleb's semantic proving style is not limited to calculus, but extends to other domains. Indeed, Caleb suggested that he employs similar strategies in abstract algebra, a domain that is less obviously visual than calculus and arguably less visual than linear algebra as well. As we did not observe Caleb writing proofs in other domains such as abstract algebra, we cannot be certain if this interpretation is correct and hence view this data as suggestive.

4.2. The influence of teaching and cognitive resources

A natural question to ask is whether, and how, Caleb's proof-writing strategies in linear algebra and calculus were influenced by how Caleb was taught these subjects. When asked about linear algebra, Caleb claimed that his syntactic strategies were a consequence of his lack of familiarity with the material.

> *Interviewer:* For the linear algebra problems, I noticed that you rarely used specific examples, specific matrices… Can you tell me a little bit more about that?
> *Caleb:* I think that it probably had a lot to do with my lack of familiarity with that material. But I think also, I don't know, calculus I like to go to the graphs. And because of my lack of familiarity [with linear algebra], one of my general proof writing strategies is, if I really have no idea, just, you know, write down all the definitions and see if you can, you know, make them meet in the

[3] Caleb began his work on the Neutral-Easy task in calculus by writing the definition of an even function.

middle. Or a lot of times just using the logic of the symbolism, you know, can get you… I've written proofs where I have no idea what's going on. Just manipulating the definitions enough and knowing that I'm being logically sound, I can make a proof that is acceptable whereas in these proofs, I understood each one so I wanted to understand them before I wrote them and I knew that wasn't going to happen on the linear algebra proofs so I wrote down the definitions and see where that would take me.

When asked to describe his linear algebra course, Caleb noted that he had taken linear algebra several years ago, and while there were proofs both in the lectures and as homework problems, the course was less proof-oriented than his advanced calculus course. He described the course as "a pretty straightforward lecture class…, pretty rote, and not very engaging… like he [the professor] was reading out of the book".

It is a natural conjecture that his advanced calculus course would emphasize a semantic approach to proving, but this was not the case.

Caleb: It's funny because my [advanced] calculus[4] class, I loved it, I thought the professor was amazing, but my professor said to me right in the beginning, every definition I put on the board, I want you to put on an index card and I want you to have all of them memorized, and at the beginning of every proof, I want you to literally just write down what you know, write down what you need to get to, and try to have them meet in the middle.

The interviewer was surprised by this response and asked Caleb for clarification.

Interviewer: So even though the professor wanted and encouraged what you were doing for the linear algebra problems, you still kind of…
Caleb: Yeah. The thing is, that strategy only came from that class. No other professor taught me that strategy. So I think that when I feel kind of lost at sea, that's the strategy that I go to immediately because I know at least something might happen, but when I do have access to the material, I like to understand it before I start writing proofs about it.

Caleb later indicated that his advanced calculus professor also gave Caleb "access to the material" in terms of graphical representations.

Caleb: For my advanced calculus class, [my professor] was just like, you know, let's draw a graph and let's really look at it. He was just very animated and engaging and explained something three different ways.
Interviewer: How would he use diagrams and examples in that class?
Caleb: He would use them very often. He would give us the definition of something, say a neighborhood, and he would draw a graph. I remember there was something really illuminating that he did. I can't remember but it was like Cauchy sequences and even if it was not a connected function and they were like getting closer like this and getting closer like this [sketches a sequence on a Cartesian plane that appear to consist of two subsequences approaching the same limit, one from above and one from below], he was just very demonstrative on the board, multiple examples. If you asked him a question, he would go to the board and draw a picture. Very graph oriented […] I do think that his mental framework was that there was proof-writing and the material. They converged, but they were like two separate things.

Our interpretation of this data is that Caleb's analysis professor enabled him to engage in semantic proof productions (Weber & Alcock, 2004). This was not done by providing Caleb with a model for writing proofs in that way; on the contrary, his professor explicitly urged Caleb to produce proofs strictly syntactically. Rather, his professor helped Caleb by helping him build rich and meaningful informal representations of the concepts he was studying and motivating his interest in the material. In other words, his professor helped Caleb build the

[4] Follow-up questions from the interviewer confirmed that when Caleb referred to his calculus professor, he was referring to the professor of his proof-oriented junior-level advanced calculus course.

cognitive resources necessary to produce proofs semantically; depending on one's theoretical perspective, we might say that the professor enabled Caleb to develop rich concept images (Tall & Vinner, 1981), semantic understandings (Weber & Alcock, 2004), or useful personal example spaces (e.g., Sinclair et al, in press). Caleb lacked this understanding and interest in linear algebra and hence was unable to approach proofs in a meaningful way to him and had to resort to writing the starting points, ending points, and trying to meet in the middle.

4.3. Time as a factor in choosing a proving strategy

In this study, Caleb was given ten minutes to work on each proving task. We chose this time allotment because we thought this would be analogous to students taking an exam in a proof-oriented task, in which students are typically asked to write four or five proofs in a 50-minute period. In his interview, Caleb indicated that he preferred semantic proving strategies because they were based on his understanding, but at several points, he said he regretted using these strategies since they might have inhibited his performance in his mathematics courses. After articulating how graphs aided his proof writing in calculus, Caleb said:

> *Caleb:* I don't know about if that is always conducive for me. Like you saw a couple of times, I'll kind of get caught up in, well I can explain it to you in two seconds if you just let me draw a graph and I think that sometimes actually hinders me rather than helps me when it comes to sitting down for an exam in math […] The way I was doing with the linear algebra proofs yesterday, I'm going to go into mechanical mode and the definitions are the definitions, I don't think I could do this with calculus proofs. And maybe I should. Maybe it would help me. I don't know […] A lot of times I have to fight it because it holds me back in a lot of cases.
> *Interviewer:* You mean in terms of time constraints?
> *Caleb:* Time constraints. Not in terms of understanding, but in terms of, here's an exam, you have this long to do it, and you have to maximize your score.

In contrast to his preferred method, when working on the Syntactic-Hard linear algebra task, Caleb was tempted to use a common student strategy that might obtain partial credit for a proof attempt, but was not conducive toward understanding. When struggling to prove that a square matrix A with the property that $A^3 = 0$ could not have a non-zero eigenvalue, Caleb said, "In a limited amount of time like this, my inclination would be to gloss over the fact that I'm not sure what's going on and I'm going to write a statement that says 'since $A^3 = 0$, then A is the zero matrix'. No, I can't really say that. That's such a broad leap."

We find it disappointing that Caleb feels the need to "fight" trying to write proofs meaningfully on his exams due to time constraints, but instead go into "mechanical mode" and regrets that he feels his failure to fight this impulse holds him back in his courses.

5. Discussion

This paper presents a case study of one student's reasoning working on proving tasks in two different domains. As with any case study, the generalizability of the findings of this case study is intrinsically limited. We begin by stating three findings that we believe are *not* generalizable and then discuss what general themes can be drawn from our data.

Caleb's lack of semantic proving strategies were attributed to his lack of familiarity with the material in linear algebra. However, not all students will abandon semantic reasoning on concepts that are new or unfamiliar to them (e.g., see Dahlberg & Housman, 1997, and Weber, Brophy, & Lin, 2008, for illustrations of how students can creatively generate examples and diagrams to understand definitions of concepts that are new to them in a short period of time). Caleb described his use of syntactic strategies as undesirable and lacking meaning, but there are other case studies of students who found such strategies meaningful (e.g., Weber, 2009). Although Caleb did not use his real analysis professor's proof constructions as a model of how he wrote proofs, other students are influenced by their professor's teaching (e.g., Weber, 2004). In summary, Caleb would not engage in semantic reasoning about concepts that he was not familiar with, found syntactic proof strategies to lack meaning, and did not mimic his real analysis professor's proof approaches. Although there are most likely other students like Caleb, we emphasize that we do not believe that all or most students behave in this way.

We do propose the following general conclusions can be drawn from the data that we present. First, Caleb exhibited different proving strategies when working on calculus and linear algebra tasks. This is an existence proof that for some students, their proving styles depend on the domain they are studying. This suggests that when researchers or teachers assign proving styles to students, it is desirable to investigate their proving strategies in more than one domain, something that has usually not been done in the mathematics education literature. Second, Caleb's case study illustrates how time pressure may influence how some students approach proving tasks and, more generally, how they seek meaning in advanced mathematics. Caleb claimed to use semantic reasoning because he valued understanding and this type of reasoning engendered understanding. However, he also lamented that he might sometimes do better abandoning these strategies to produce proofs in a less meaningful but more efficient manner. As an extreme example, Caleb noted that on exams, he would sometimes engage in what might be described as "wishful thinking"—conjecturing a property that would make his proof easier to hide the fact that he did not understand what was happening. Third, Caleb's case illustrates that students' proving styles are not solely a function of the student but *can sometimes* be the result of instruction. Although Caleb's advanced calculus professor explicitly encouraged Caleb to use a syntactic reasoning strategy—advice that Caleb generally ignored (except, sometimes, when he was less comfortable in the domain)—his teaching nonetheless influenced Caleb. By providing Caleb with opportunities to develop rich graphically-based understandings of concepts in real analysis and fostering his interest in the subject, he enabled and motivated Caleb to generate informal graphical explanations prior to proving theorems and using these explanations as a basis for the proof that he constructed.

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. & Inglis, M. (2009). Representation systems and undergraduate proof production: A comment on Weber. *Journal of Mathematical Behavior 28*, 209-211.

Alcock, L. & Simpson, A. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 57*, 1-32.

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*, 77-110.

Alcock, L. & Weber, K. (2010). Referential and syntactic approaches to proof: Case studies from a transition-to-proof course. In *Research in Collegiate Mathematics Education VII* (pp. 93-114).

Burton, L. (2004). *Mathematicians as Enquirers: Learning about Learning Mathematics.* Berlin: Springer.

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: A systematic and critical review.* London: Learning & Skills Research Centre.

Dahlberg, R., & Housman, D. L. (1997). Facilitating learning evants through example generation. *Educational Studies in Mathematics, 33,* 283-299.
Duffin, J. & Simpson, A. (2006). The transition to independent graduate studies in mathematics. *Issues in Mathematics Education, 13*, 233-246.
Entwistle NJ (1998). Improving teaching through research on student learning. In J. J. F. Forrest (Ed.) *University teaching: International perspectives*. New York: Garland.
Moutsios-Rentzos, A. (2009). Styles and strategies in exam-type questions. In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education*. Thessaloniki, Greece: PME.
Pinto, M., & Tall, D. (1999). Student constructions of formal theory: Giving and extracting meaning. In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 65–73). Haifa, Israel: PME.
Pinto, M., & Tall, D. (2001). Following students' development in a traditional university classroom. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 57–64). Utrecht, Holland: PME.
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.
Poincare, H. (1913). *The foundations of science.* (translated by G.B. Halsted). New York: The Science Press.
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.
Riding, R., & Cheema, I. (1991). Cognitive styles: An overview and integration. *Educational Psychology, 11*(3-4), 193-215.
Sinclair, N., Watson, A., Zazkis, R., & Mason, J. (2011). The structuring of personal example spaces. *Journal of Mathematical Behavior*, 30, 291-303.
Tall, D. & Vinner S. (1981). Concept image and concept definition in mathematics, with special reference to limits and continuity. *Educational Studies in Mathematics*, 12, 151-169.
Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161-177.
Vinner, S. (1991). The role of definitions in teaching and learning. In. D. Tall (Ed.) *Advanced mathematical thinking*. Dordrecht: Kluwer.
Weber, K. (2004). Traditional instruction in advanced mathematics: 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. (2009). How syntactic reasoners can develop understanding, evaluate conjectures, and construct counterexamples in advanced mathematics. *Journal of Mathematical Behavior, 28*, 200-208.
Weber, K. & Alcock, L. J. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56* (3), 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., Brophy, A., & Lin, K. (2008). Learning advanced mathematical concepts by reading text. *Proceedings of the 11th Conference on Research in Undergraduate*

Mathematics Education. Retreived 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. *Journal of Mathematical Behavior, 28,* 212-216.
Weyl, H. (1940). The mathematical way of thinking. *Science*, 92, 437-446.

Appendix
Tasks, in the order presented to Caleb.
Session 1: Linear Algebra

Neutral-Easy: Prove that $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : xyz \geq 0 \right\}$ is not a subspace of $\mathbf{R}^3$.

Syntactic-Hard: Suppose that A is a square matrix for which $A^3 = 0$, and r is any non-zero real value. Prove that r is not an eigenvalue of A.

Syntactic-Medium: Suppose $\{v_1, v_2, \ldots, v_n\}$ forms a basis for a vector space V. Prove that $\{v_1, v_1 + v_2, \ldots, v_1 + v_2 + \cdots + v_n\}$ also forms a basis for V.

Semantic-Medium: Let $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ be unit vectors in $\mathbf{R}^3$ with $|u_3| > |v_3|$. Prove that the projection of u onto the x-y plane is shorter than the projection of v onto the x-y plane.

Semantic-Hard: Given that $U = \left\{ \begin{bmatrix} a - 2b + 5c \\ 2a + 5b - 8c \\ -a - 4b + 7c \\ 3a + b + c \end{bmatrix} : a, b, c \in R \right\}$ is a subspace of $\mathbf{R}^4$, prove that $\dim(U) = 2$.

Neutral-Hard: Suppose T is a non-invertible 2×2 matrix. Prove that there exists a nonzero 2×2 matrix S such that TS is the zero matrix.

Neutral Medium: Suppose U and W are subspaces of $\mathbf{R}^4$, dim(U)=2 and dim(W)=3. Prove that the intersection of U and W must contain more than one vector.

Session 2: Calculus

Neutral Medium: Suppose $f(0) = f'(0) = 1$. Suppose $f''(x) > 0$ for all positive x. Prove that $f(2) > 2$.

Semantic Hard: Prove that the only real solution to the equation $x^3 + 5x = 3x^2 + \sin x$ is x=0.

Neutral Easy: Suppose $f(x)$ is a differentiable even function. Prove that $f'(x)$ is an odd function.

Syntactic Medium: Prove that $a^2 + ab + b^2 \geq 0$ for all real numbers a and b.

Neutral Hard: Suppose $f''(x) > 0$ for all real numbers x. Suppose a and b are real numbers with $a < b$. Define $g(x)$ as the line through the points $(a, f(a))$ and $(b, f(b))$. Prove that for all $x \in [a,b]$, $f(x) \leq g(x)$.

Semantic Medium: Prove that $\int_{-a}^{a} \sin^3(x)\,dx = 0$ for any real number a.

Syntactic Hard: Let f be differentiable on $[0,1]$, and suppose that $f(0) = 0$ and f' is increasing on $[0,1]$. Prove that $g(x) = \dfrac{f(x)}{x}$ is increasing on $(0,1)$.

INVESTIGATING TEACHING PRACTICES WHEN PRESENTING PROOFS: THE USE OF EXAMPLES

Melissa Mills
Oklahoma State University

This study combines interview data and observation data to investigate the teaching practices of mathematics faculty members when teaching upper-division proof-based undergraduate mathematics courses. Four case studies of faculty members at a large research institution who were teaching in different mathematics content areas are used to construct a model describing the ways in which examples are used to motivate and support proof presentations in class.

Key words: model of example usage, proof presentations, examples, undergraduate teaching practices

1. Introduction

Although there have been many calls for studies addressing the teaching practices of university teachers, the literature contains very few responses (Harel & Fuller, 2009; Harel & Sowder, 2007; Speer, Smith, & Horvath, 2010). In particular, there has been very little research addressing the teaching practices of faculty members in upper-division proof based courses (Weber, 2004). In fact, after an extensive search of the literature, Mejia-Ramos and Inglis (2009) found 131 research papers addressing writing, reading, and understanding of proof for undergraduates, but none of these papers described how proofs were presented by instructors in class.

One of the main purposes for presenting proofs in class is that the instructor, as an expert, models the proving process (Fukawa-Connelly, 2010). Examples are often used by successful mathematicians and advanced mathematics students to help them better understand the concepts behind a proof (Inglis, Mejia-Ramos, & Simpson, 2007; Weber, 2011). Thus, when instructors are modeling the mathematical behavior of proof writing, they may use examples in their presentations in the same ways that experts would. The goal of this study is to investigate and build a model to describe ways in which faculty members use examples when presenting proofs in courses in undergraduate proof-based mathematics courses. Interview data and video data from four different faculty members teaching abstract algebra, analysis, number theory, and geometry are analyzed to determine the ways these instructors use examples to motivate and support their presentations of proofs in class. The main result of this paper is a model of example usage in proof presentations across different content areas. It should be noted that this study is a part of a larger study which examines several different aspects of proof presentations, including example usage, student interaction, and methods of fostering students' strategic thinking.

2. Research Questions

In what ways are examples used in class to motivate and support statements of claims and presentations of proofs in an upper-division proof-based mathematics course? What is the pedagogical motivation of the instructor for the use of particular examples in proof presentations?

3. Literature Review

At the collegiate level, there are few studies focusing on teaching practice, i.e. "what teachers do in and out of the classroom on a daily basis" (Speer et al., 2010). A foundational understanding of teaching practice contributes to further understanding of the phenomenon of teaching and learning. In particular, there is value in focusing in on small, meaningful aspects of practice that mathematicians already use in the classroom (Speer, 2008).

Example usage in mathematics classrooms is one important aspect of teaching practice that has been examined in the literature (Fukawa-Connelly, Newton, & Shrey, in press; Watson & Mason, 2005; Watson & Shipman, 2008; Zodik & Zaslavsky, 2008). Some believe that the goal in teaching advanced mathematics is to help students to replace the use of examples and empirical reasoning with formal, generalized proof schemes (Harel & Sowder, 1998). On the other hand, even advanced mathematics doctoral students use exploration of examples as part of the process for constructing an original proof (Alcock & Inglis, 2008), and working mathematicians sometimes generate examples to help them read and comprehend a mathematical argument (Weber, 2011). Therefore, it appears that examples are used by successful mathematicians and graduate students to help them understand and construct proofs. In fact, the ability to apply a proof strategy to a specific example is listed among the dimensions of proof comprehension (Mejia-Ramos, Fuller, Weber, Rhoads, & Samkoff, 2012). They claim that "comprehending a proof often involves understanding how the proof relates to and can be illustrated by specific examples—that is, being able to follow a sequence of inferences in the proof in terms of a specific example," (p. 14). Thus, if the goal of presenting proofs in class is to "model the mathematical behaviors" of successful mathematicians, examples should play a pivotal role (Fukawa-Connelly, 2010).

Examples serve as an important bridge between concrete computation and the abstraction required to construct a general argument. Therefore, instructors should not encourage students to abandon empirical reasoning entirely but to appropriately pair intuitive examples with deductive arguments (Inglis, Mejia-Ramos, & Simpson, 2007). This study will examine ways in which instructors of advanced mathematics pair examples with their presentations of proofs in class.

Attending to proof presentations in class is one of the primary ways in which students construct their understanding of what constitutes a proof (Weber, 2004). There is evidence to suggest that when teaching proof-based courses, some professors spend large portions of their class time (between one third and two thirds) presenting proofs (Mills, 2011). Several recent studies have used faculty interviews to investigate the pedagogical views of faculty members concerning proof presentations in class (Weber, 2011; Yopp, 2011; Alcock, 2009; Harel & Sowder, 2009; Hemmi, 2010; Lai, Weber, & Mejia-Ramos, in press). Some of these studies discussed a relationship between proof presentations and the use of examples. Several instructors mentioned that they often accompany a proof with an example (Weber, 2011). Alcock (2009) identified 'instantiation of definitions and claims' as one of the four proof-related skills that instructors are trying to teach. Observations of a particular professor throughout the course of a semester revealed that he modeled the mathematical behavior of ‘example exploration and generalization’ when presenting lectures in class (Fukawa-Connelly, 2010).

One contribution that the present study makes to the literature is that it combines faculty interviews with observation data to investigate what faculty members think about the pedagogy of examples and proof presentations as well as catalog their actual behaviors in the classroom. This will allow for the investigation of their teaching practices, which is more in line with the type of studies called for by Speer et al. (2010). This study will present a model of instructors’ example usage in proof presentations that is grounded in the observation and interview data and also consistent with the literature.

In order to investigate instructors’ use of examples, it is crucial to define what exactly is meant by the word “example.” Mason & Watson (2005) define an example as “anything from which the learner may generalize” (p. 3). This implies that there is some particular element in an example that could be generalized by the learner. This broad definition includes illustrations of concepts, placeholders for general definitions or theorems, solutions demonstrating a technique, or applications to motivate the mathematics. The next few paragraphs outline some of the different uses of examples found in the literature.

3.1 Start-Up Examples and Pattern Generalization

When introducing a new concept, *start-up examples* can be used to motivate basic intuitions or claims. These examples are easily accessible to the learner and help to prepare them for a more difficult concept (Michner, 1978). Another class of examples that occur before the presentation of new content are examples that *generalize a pattern*. In their description of proof schemes related to mathematical induction, Harel (2001) described students’ use of *pattern generalization* in two ways. One way students used pattern generalization is as an empirical proof scheme, showing that a property held for a few “randomly selected” numbers. The second is the use of pattern generalization to extract a process that can generalize to a deductive proof. It may be that instructors can use pattern generalization to help students prepare for the statement of a theorem or to help them see the need for a more general proof (Rowland & Bills, 1999).

3.2 Boundary Examples

Examples are often used when presenting a definition of a new concept. What characterizes an example as a *boundary example* is that it helps the students distinguish between having and not having a specified property (Mason & Watson, 2001). These examples can help the students to more clearly define the edges around a certain definition. Historically, examples have caused mathematicians to adapt definitions and to gain a better understanding of the concepts. Latakos’ (1976) *Proofs and Refutations* gives a detailed classroom vignette of students who are grappling with a proof. The generation of problematic examples helps the class to better solidify their understanding of the definitions involved in the construction of the proof.

3.3 Instantiation

Examples can also be used to *instantiate* claims or definitions (Alcock, 2009). These examples serve to give the students concrete instances or applications of a claim or definition. A (dynamic or static) picture may also serve as an instantiation of a claim or definition (Mason & Watson, 2005). Fukawa-Connelly, Newton, & Shrey (in press) focused on the use of examples in a proof-based course by describing in detail how a faculty member used examples to instantiate the definition of a mathematical group in an abstract algebra class.

3.4 Generic Examples

Generic examples are carefully selected examples that can enable students to see the general arguments of a proof embedded in the particular example (Rowland, 2002). The general proof and the example may be worked side-by-side so that the students can extract the general method of the proof from the example, in the same way that some mathematicians read proofs on their own (Weber, 2011). Michner (1978) uses the term *model examples* to mean paradigmatic, generic examples that are indicative of the general case. While it is unclear if Michner was talking about examples that could be used in a proof presentation, the idea of an example that has the properties of the general case is similar. An instructor may also present a picture representing the general case side-by-side with a proof, as observed by Weber (2004).

4. Methodology

Faculty members at a large comprehensive research university who were teaching proof-based upper-division mathematics courses during between August 2010 and August 2011 were asked to participate in the study. Four experienced, tenured, faculty members agreed to participate in a one hour interview and agreed to allow their lectures to be video-taped approximately once every two weeks throughout the semester. They all taught in a lecture style, with the instructor primarily teaching at the board while the students were listening, taking notes, and sometimes answering questions and participating in class discussions. Throughout this paper, all participants will be referred to using masculine pronouns regardless of their gender.

Interviews were transcribed and analyzed using the constant comparative method (Glaser & Strauss, 1967) to establish codes pertaining to the pedagogical views of the participant. In the interview, the participants were asked to describe what they do to help the students understand a proof that they present in class. Though the participants were not specifically asked about the ways in which they used examples to support their proof presentations, all participants mentioned examples in their interview in some way. For this study, sections from the data that had codes pertaining to example usage were selected and set a backdrop for the instructors' observed usage of examples from the observation data.

The analysis of the video data occurred in several phases, using the grounded theory approach (Glaser & Strauss, 1967). First, I viewed the videos and took notes on what was happening in each time interval. Then all of the instances of proof presentation in the observation data were transcribed. For this study, I have identified all of the instances in the data when examples are used to support the proof of a claim.

When analyzing the interaction between examples and proof presentations, one of the ways in which I determined the pedagogical intention of the example was the timing of the example presentation. Examples presented before the statement of a claim serve to prepare the students for the statement of the claim, or may call their attention to a pattern, or give insight into the conditions of the claim that will be presented. Examples presented after the presentation of a claim but before the proof serve to help students understand some aspect of the statement of the claim, or may help prepare the students for the presentation of the proof. Examples presented during the proof may be intended to highlight some aspect of the proof structure, or to explain or remind students of a particular concept associated with the proof. Examples presented after the proof may instantiate the claim or apply the proof method to an example.

Since the instructors are teaching different content areas, it makes sense to consider these as separate but interrelated case studies. Although data from the different cases will be combined in the model, no attempt is made to evaluate the methods used by the participants. Rather, the goal of this study is to describe and model the use of examples when presenting proofs across mathematical content areas.

5. Results

Because the four faculty members were teaching different subject matter, the data is presented in the form of four interrelated case studies (Zodik & Zaslavsky, 2008). The setting of each case study is described, including some excerpts from the interviews that relate to example usage. Then a description of the types of examples is presented that incorporates the uses of examples from all participants. Excerpts from the observation data are used to illustrate the different aspects of the example types.

5.1 Settings

5.1.1 Introduction to Modern Algebra. Introduction to Modern Algebra is a junior-level course that is required for all mathematics and mathematics education majors. It serves as both an introduction to proof and an investigation of modern algebra. This course is a prerequisite for every other proof-based course. According to the university course catalog, it covers an introduction to set theory and logic, elementary properties of rings, integral domains, fields, and groups. In the semester that Dr. A taught the course, he used Durbin's (2009) *Modern Algebra: An Introduction*. The observation data showed that he used the textbook both to organize the presentation of the material and to assign homework problems. The class enrollment consisted of 24 students with a diverse range of majors: six math education majors, eight math majors, six engineering majors, two computer science majors, one geography major, and one chemistry major. There was one sophomore, and the rest were split almost evenly between juniors and seniors.

Initial analysis of the observations showed that Dr. A spent approximately 40% of his class time presenting proofs (Mills, 2011). Dr. A used examples in some way in 13 out of the 16 proofs captured on video.

In his interview, Dr. A didn't talk about the use of examples specifically, but he did mention that he is very picture oriented and typically draws pictures. He also said, "Another thing I find helpful is if you have organized the proof in such a way that also with the steps you have kind of a continuous picture that you are filling in at the same time." He may have been alluding to a picture that would be worked side-by side with the general proof (Weber, 2004). Unfortunately, the observation data did not capture Dr. A using this type of picture, though he did use examples in his presentations in several different ways.

5.1.2. Geometry. Dr. G was teaching the senior level course in Geometry required of all math education majors. The university catalog describes this course as an axiomatic development of Euclidean and non-Euclidean geometries. The textbook used was Venema's (2002) *Foundations of Geometry*, but the observation data showed that the structure of the lectures did not follow the book. The order of presentation of the content was completely different than the order in the book, and some of the named theorems in the lecture had slightly different names in the book. Also, in the observation data, Dr. G never referenced the book. In a recent informal conversation, Dr. G confirmed that he did not follow the book but expected the students to use the book as a reference. Additionally, he mentioned that the homework problems were a mixture of those he wrote himself and those available in the textbook. The class enrollment consisted of 9 students: four math education majors, four math majors, and one engineering major. Eight of the students were seniors, and one was a junior.

Initial analysis of the six video observations showed that Dr. G used approximately 70% of his class time on presenting proofs (Mills, 2011). In the 22 proofs observed, Dr. G used examples in 19 of them.

In his interview, Dr. G spoke about how he would present the proof of the Saccheri-Legendre Theorem, which states that in neutral geometry (Euclid's first four axioms), the angle sum of a triangle is less than or equal to 180 degrees. He said he would lead up to the theorem by giving examples of simple geometries in which the angle sum is less than 180 degrees, and added that he typically leads up to big theorems with examples. He also mentioned that he draws pictures, because "it's geometry," but he did not elaborate on the pedagogical purposes of the pictures that he used. Excerpts from the observation data shed more light on the possible purposes that the pictures served in Dr. G's proof presentations.

5.1.3 Number Theory. Dr. N taught the senior level Number Theory course, which is required of all math education majors. The university course catalog states that this course covers divisibility of integers, congruences, quadratic residues, distribution of primes, continued fractions, and the theory of ideals. He used the textbook *Elementary Number Theory and its Applications* by Rosen (2011). The observation data showed that he followed the basic outline of the textbook, often referring to section numbers in class, and he also used the book for homework problems. There were 14 total students in the class: seven math education majors, six math majors, and one engineering major. Thirteen of the students were seniors, and there was one junior.

In the video observation data, Dr. N spent approximately 35% of class time presenting proofs (Mills, 2011). He used examples in 7 of the 9 proofs in the observations.

In his interview, Dr. N mentioned that he likes to do computation before stating the claim that he is going to prove. He said, "Well, if it's a proof of a pattern, then I certainly emphasize computation. First, you have to compute a lot to try to figure out what the pattern is. And, uh, so, you should always do some computation before any proof. You know, or some idea of why you are going to go into the proof." These examples allow the students to discover the statement of the proof from computations. This is called *pattern generalization* in the literature (Rowland & Bills, 1999; Harel, 2001; Inglis, Mejia-Ramos, and Simpson, 2007).

Another usage of examples was noted by Dr. N in his interview. He talked about developing the skill of "looking at theorems and trying to understand how to produce examples out of it. So, I spend a lot of time doing that. So, here's a theorem, can you give me like a special case of it?" Dr. N called this the technique of specialization, referencing Polya's (1973) famous work on problem solving. The observation data does show one instance when Dr. N used this technique in class.

5.1.4 Introduction to Modern Analysis. Dr. C taught a senior level course in modern analysis. According to the university course catalog, the course covers properties of the real numbers, sequences and series, limits, continuity, differentiation and integration. Dr. C used the book, *Analysis with an Introduction to Proof* (Lay, 2005). The video observations showed that Dr. C used the book regularly in class, referencing numbered theorems from the book, reading definitions straight from the book in class, and sometimes even projecting pages from the book onto the board. There were 9 students in the class: five math education majors, two math majors, and two engineering majors. Eight were seniors, and one was a junior.

Initial analysis of the video observation data showed that Dr. C used 49% of the observed class time presenting proofs, and that examples were used in 11 of the 22 proofs observed. Of the four participants, Dr. C's example usage varied the most.

In the interview, Dr. C spoke of using pictures in proofs, and how students interpret pictures. He said, "Sometimes the algebra books talk about functions in a very abstract way. And they'll draw like a set here, and an arrow to a set here. I'm not clear that students have any sense in how to process that kind of a thing. Uh, maybe… but, I don't know. Um, I even think that with graphs of functions there's something to be done. I think that the picture itself, you can't always be sure that students process a picture in the way that you want them to." Even though he questioned whether students understand these types of pictures, the observation data showed Dr. C did use such pictures of sets and mappings in one of his proofs.

Dr. C talked about how he uses examples in his own work to help him make sense of a statement of a claim. He said "I have to play with that statement in my brain to make sense of it. Well, how do I make sense of it? I start looking at examples." Because Dr. C uses examples in this way in his own work, he said that he uses examples and computation in class to "prepare the students' minds" for the statement of a theorem. In the observation data, Dr. C used this method several times when presenting proofs.

5.2 Description and Analysis of the Types of Examples Used in Proof Presentations

Upon analysis of the proofs that were captured in the observations, three primary purposes for examples surfaced: to motivate or support the statement of the claim, to motivate or support the proof, or to reinforce the mathematical content underlying either the claim or the proof. The timing of the examples in relation to the statement of the claim or the presentation of the proof often sheds light on the pedagogical purpose that the example serves. Initial interviews with the participants also contain some comments about the participants' pedagogical thoughts concerning examples. While it may be the case that planned examples and spontaneous examples serve slightly different pedagogical purposes, I do not make this distinction in this analysis (Zodik & Zaslavsky, 2008).

In the next few sections, I will describe in detail the different types of examples that I observed in my data, including excerpts from the observation data to support my claims.

5.2.1. Warm-up Examples. These examples are used to motivate basic intuitions or claims (Michner, 1978), or to "prepare the students' minds" for the statement of the claim. This type of example usually occurs before the statement of the claim. One instance of this type of example appeared in Dr. C's analysis lecture.

Dr. C began with two open intervals in **R**, (0,2) and (1,3), noting that the union of these sets is still open. Then he asked, "What about an infinite collection of open sets?" He created the collection of sets $O_j = (0,2-\frac{1}{j})$, and $C_j = [0,2-\frac{1}{j}]$. Then, prompted by the instructor, the class discussed whether or not 2 is included in $\bigcup_{j=1}^{\infty} O_j$ or $\bigcup_{j=1}^{\infty} C_j$. After the warm-up example, Dr. C presented and proved the theorem: "(a) The union of any collection of open sets is open. (b) The intersection of any finite collection of open sets is open." After the proof of that theorem, he presented the corresponding theorem regarding closed sets with no general proof.

5.2.2. Pattern Generalization. Examples showing a pattern that are used before the statement of the claim to provide students with intuition about the claim will be called *pattern generalization examples.*

In his interview, Dr. N claimed that he sometimes uses this technique to construct theorems: "You're going from examples to theorems… you go through a lot of examples, you try to find something that's always true, and then you conjecture a theorem." Dr. N used this technique in class when he was lecturing about estimating the number of steps in Euclid's Algorithm, an algorithm used to find the greatest common divisor of any two integers *a* and *b*.

He began by stating that the "worst cases" for Euclid's Algorithm (meaning pairs of numbers that will maximize the number of steps) happen when the two numbers are consecutive Fibonacci numbers. He illustrated Euclid's Algorithm for a=13 and b=8 to show that there are 5 steps. He then used a=144 and b=89 to show that there are 10 steps. Next, he stated Lamé's Theorem: "The number of steps in Euclid's algorithm for integers *a* and *b* is less than or equal to five times the number of digits of the smallest of *a* or *b*." Because the pattern in the computational examples foreshadowed the statement of the theorem, this was coded as *pattern generalization.* He proceeded to give a general proof of Lamé's Theorem.

5.2.3. Critical Examples. Critical examples serve to highlight the necessity of the hypotheses of the claim. I chose the word "critical" because it can mean both "analytical judgment" or "providing textual variants." Critical examples can occur before or after the statement of the claim.

Dr. C talked about critical examples in his interview: "I think through computing examples and looking at cases where the theorem does and does not hold, I think you can prepare them for understanding the parts of the hypotheses." The observation data showed Dr. C using some *critical examples* to highlight the necessity of the hypotheses for the Heine-Borel Theorem.

One day, Dr. C proved that the set {1,3,5,7…} is not compact in **R** using the definition. Then, he presented the Heine-Borel Theorem, which states that a closed, bounded set of real numbers is compact. He asked the students why the previous set was not compact, and the students said "It is not bounded above." Then the class proceeded into an instructor-led discussion of whether or not the set was closed, using the definition of limit points and a sketch of the set on a number line. Dr. C then said, "Let's have another example [of a non-compact set]. What property of the Heine-Borel theorem should we contradict now?" The students said that they need an open set (note that they really only needed a set that was not closed). Dr. C then asked for an open set, and with some prodding, students chose (0,1). Dr. C then proved, using the definition, that (0,1) is not compact. These two examples showed the necessity of the closed and bounded conditions in the Heine-Borel theorem.

5.2.4. Instantiation of the Claim. When an example is used after the statement of the claim to give an instance when the claim holds, this is called *instantiation of the claim* (Alcock, 2009). All of the participants used examples to instantiate the claim at different times in their lectures. They all occurred after the presentation of the claim and may or may not have been followed by a deductive proof. There was one instance where instantiation occurred after the proof. Instantiation of a claim may support the statement of the claim or prepare the students for the proof, depending on the situation.

Dr. N presented the claim that "The GCD of two Fibonacci numbers f_n and f_m is $f_{\gcd(n,m)}$." He chooses f_8 and f_{12}, stating that *n* and *m* do not have to be "next to each other." He then says, "f_8 turns out to be, uh, 21? Is that right? And f_{12} is 144, and the greatest common divisor of that is 3, which is equal to f_4. So, that's an example of that. So, um, that's an amazing fact about the Fibonacci numbers." He then proceeds to give a start to a proof. He does not present a complete proof, leaving the remaining details as a homework problem.

Dr. G used a pictorial example to instantiate a claim. He had just written this statement on the board: "In a projective plane where each line meets *n* points, (a.) there are a total of $n^2 - n + 1$ points; (b.) there are a total of $n^2 - n + 1$ lines." He then said, "Before we prove this theorem, let's just talk about these projective planes a little bit. Uh, what's the smallest projective plane I can possibly have? By smallest, I mean smallest number of points and smallest number of lines." With a little bit of input from the students, he concluded that the case when $n = 3$ is the smallest projective plane, with seven points and seven lines. He drew a model of that particular projective plane, and then proceeded into a proof of the theorem. The picture served to support the students' understanding of the statement of the claim, not to support the construction of the proof. Therefore we call this a *pictorial instantiation of a claim.*

5.2.5. Generic Examples. Another type of example that was used is a *generic example* (Rowland, 2002). These examples are used in conjunction with a general proof. They must be carefully chosen to be not too easy or too difficult, but to mirror the structure of the general proof. It is interesting that the data showed a *generic example* occurring in the Number Theory class, because that is precisely the context in which Rowland (2002) described this type of example. There was no evidence of the usage of a generic example in any of the other participants' lectures, although I believe that they could occur in other mathematics content areas.

On one day of class, Dr. N gave general proof of the claim that there is a unique representation of any integer in a particular base. The theorem is stated, "Let *b* be an integer greater than one. Every integer *n* greater than or equal to one has a unique representation as $n = a_k b^k + a_{k-1}b^{k-1} + ... + a_1 b + a_0$, where $0 \le a_j < b$ for all *j*, and $a_k \ne 0$." In the middle of the proof, Dr. N stopped to give an example. The computational example mirrored the general proof, which served to help the students understand the process that they were using in general. The following excerpt occurred after Dr. N had stated the theorem and started the general proof.

> Dr. N: Ok, so I just used the same thing again, I repeat again… I'm going to continually keep going like that… And, so we repeat this process, and so 'Repeating at the *j*th step, we have $q_j = q_{j+1}b + a_{j+1}$, where $0 \le a_{j+1} < b$.' The remainder always satisfies that it's between zero and *b*. Ok, so I'm just going to keep dividing over and over again. So, if you do it in practice… it might actually be worth throwing some numbers up so that you can compare that against the theory, the theoretical formula. So, uh, '73 base 5'. Last class we said that $73 = 14 \cdot 5 + 3$, and this was a_0 *(labels the 3)* and this was q_0 *(labels the 14)*. So, then $14 = 2 \cdot 5 + 4$, and this is a_1 *(labels the 4)* and this is q_1 *(labels the 2)*. And then, um, the next quotient is $2 = 0 \cdot 5 + 2$, so this would be a_2, and this would be q_2. So... well, you can kind of imagine what the three steps are in there… and so now, do you notice anything about this process here that is allowing me to stop? What is it about this process that is allowing me to stop and say, 'I can finally stop dividing'?
>
> **Student 1: You get zero for…**
>
> Dr. N: I get zero for the quotient, Ok, that's right. So, that's the tip off, I get zero for the quotient. But, why are you forced to get a quotient that is zero? You have to detect a pattern in order to…
>
> **Student 2: The quotients are always deceasing.**
>
> Dr. N: The quotients are always decreasing. That's right, exactly right. If you look at the pattern, here, the quotients go from 14 to 2 to 0. And, whenever you do that the quotients are always decreasing. Ok? *(continues with the general proof)*

So, Dr. N used the numerical example to help the students to see how they can write the general proof, because the example showed the students that the quotients were decreasing. He also tied the general notation of the proof to the numbers in the example, so that the students could have a more concrete understanding of the general proof. In this way, working the numerical example to the side aided in the construction of the general proof.

In Dr. G's geometry class, he usually wrote the statement of the theorem on the board and reserved a space on the board for drawing pictures side by side with the general proof. Then he would fill in the picture and write the proof simultaneously, using the picture to guide the next line of the proof. Dr. G used this strategy in 19 of the 22 proofs that he presented in the observation data. Since he was drawing the general picture and writing the proof "side-by-side," these were classified as *pictorial generic examples.* Note that a *pictorial generic example* is a picture representing the general case that is used to guide a deductive proof, whereas a *generic example* is a more particular and sometimes numerical example that mirrors the proof.

5.2.6. Instantiation of a Sub-Claim. These are examples that are used during the proof of a claim to help students understand a sub-claim.

In the middle of the proof of Lamé's theorem, Dr. N claimed that he can use the log base ten function to determine the number of digits of a number. The class seemed puzzled by this notion, so Dr. N presented an example to *instantiate the sub-claim*. He gave the example of 5643, where $10^3 < 5643 < 10^4$, so therefore $3 < \log_{10} 5643 < 4$, and then stated that the number of digits of the number 5643 is the ceiling of $\log_{10} 5643$.

5.2.7. Application of the Proof Method. In their model for proof comprehension, Mejia-Ramos et al. (2012) claim that applying the proof method to an example would give evidence of proof comprehension. Although this did not occur in my observation data, I still believe it is a type of example that could be used after a proof presentation to show the importance of a proof technique.

5.2.8. Instantiation of Definitions, Concepts, or Notation. These examples serve to reinforce the mathematics content underlying the claim or the proof. They may occur at any time, and sometimes appear to be spontaneously generated by the instructor.

When proving a statement about bounded sequences, Dr. C asked the class to give some examples of a bounded sequence (they had previously defined a bounded sequence). These examples served to *instantiate the definition* of a bounded sequence. One student suggested "1, 2, 3, 3, 3, 3, 3…" Dr. C said that this is a bounded convergent sequence, and asked if the students if they could think of a bounded sequence that does not converge. Another student suggested "$(-1)^n$." Then Dr. C asked the class what it means to say that a general sequence is bounded, and the students began to reconstruct the definition. He then proceeded to prove the general claim using this definition.

On one class day, Dr. A presented a set and a relation on that set, and proved that it was an equivalence relation. In the proof, he asked the students to tell him some of the elements in an equivalence class of a particular element. Rather than an instantiation of the claim itself, this is an *instantiation of the concepts* involved in the proof.

Instantiation of notation is also used by Dr. C. When proving that $s_n = \dfrac{n+1}{n+2}$ converges, he first asked students, "If n=100, what is s_n?" Then he asked them to compute s_1, s_2, and s_3. These examples were used to help the students make sense of the notation and begin to see the pattern of the sequence.

5.2.9. Metaphorical Examples. Whenever an example is used to compare one mathematical structure to a different (more familiar) mathematical structure, I refer to this as a *metaphorical example.* This is different from instantiation, because the instructor is comparing two different structures rather than giving an instance of a structure. Metaphorical examples may occur at any time throughout the presentation of the claim or proof, and serve to reinforce the mathematics content underlying the claim or proof.

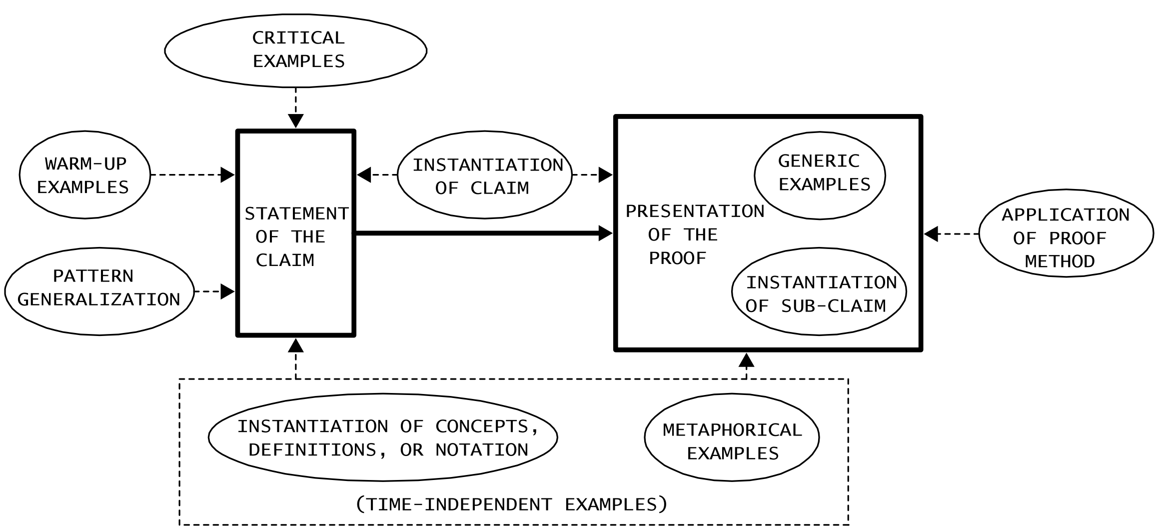
When enumerating the different types of elements in S_4, the students tried to count the four one-cycles as different elements. Dr. A wanted to discourage this behavior and help the students to understand that though the elements are written differently, they represent the same element of S_4. Dr. A said, "Well, they're all the same, so I just do that *(writes '1-cycles: (1)')*. That's all the 1-cycles we have. Yeah, you can come up and write (4), but that's equal to that, so, I just, you wrote it different. It's not different. That's the only one cycle." Then a student asks, "So, there's only one 1-cycle?" to which Dr. A replied, "It's the identity. And, your eyes are telling me that that confuses you. Ok, like we do a lot, you can write ½ equal to $^{30}/_{60}$, that's not two fractions, that's one fraction."

Since there are mathematical differences in the two structures, this is not really instantiation. Dr. A is comparing the unfamiliar structure to a more familiar structure using a metaphor, because in both instances an element can be written in different ways.

6. Model for Example Usage in Proof Presentations

The model in Figure 1 below illustrates the ways that examples were used, and how they interact with the statement of the claim and the presentation of the proof.

Figure 1:
Model for Example Usage in Proof Presentations



The diagram is a timeline moving from left to right. The placement of the ovals represents the sequence in which the examples usually occur. Critical examples could occur before or after the statement of the claim, so the oval is elongated on both sides of the claim. Dotted arrows show whether the example is used to motivate and support the claim or the proof. The examples in the dotted box at the bottom could occur at any time, so they are not sequentially listed with the rest of the examples.

- *Warm-Up Examples* occur before the statement of the claim and serve to prepare the students minds for the claim.
- *Pattern Generalization Examples* also occur before the statement of the claim and help the students to generalize the statement of the claim from concrete examples.
- *Critical Examples* serve to highlight the necessity of the hypotheses of the claim and may occur before or after the statement of the claim.
- Examples that *Instantiate the Claim* occur after the statement of the claim, and may serve to help students understand the claim, or to prepare them for the presentation of the proof.
- *Generic Examples* occur during the proof, and are written side-by-side with the proof so that students can take aspects of the particular example and apply it to the general proof.
- Examples that *Instantiate a Sub-Claim* generally occur during a proof, and support the students' understanding of a sub-claim.
- *Application of the Proof Method* can occur after the presentation of the proof, and serves to illustrate the usefulness of the proof method.
- Examples used to *Instantiate concepts, definitions, or notation* serve to reinforce the mathematics content underlying the claim or proof, and may occur at any time.
- *Metaphorical Examples* can also be used at any time. These occur when an instructor compares some aspect of a mathematical structure to a different, more familiar, mathematical structure via metaphor.

7. Discussion

The four participants were not explicitly asked about the pedagogical purposes of their example uses in the interviews, and though they did give some insight into this issue, they may have responded differently if asked directly. The observations occurred periodically, and so they represent only a snapshot of the participants' teaching. Thus, they may not be reflective of their practice in general. Despite the small size of the study and limited number of observations, I have still obtained a rich collection of examples that leads nicely to the construction of a model for example usage in proof presentations.

This study contributes to the current research in several ways. First of all, it investigates the teaching practices of university professors by combining interviews about their pedagogical thoughts concerning proof presentations with observations of their actual practice when instructing students. Secondly, it builds upon the example literature by providing empirical evidence of instructors' example usage. Thirdly, the treatment of pictures in this analysis is based on their pedagogical purposes leading to their classification into several different categories, whereas past research has lumped pictorial examples into one category. Lastly, I have identified two new types of examples that are used in proof presentations: critical examples and metaphorical examples.

The primary contribution of this paper is the presentation of a model for example usage in proof presentations. This model may be used by researchers to design studies linking example usage in proof presentations to student learning or to inform research on examples or proof presentations. It also can support instructors as they design lectures in which they present proofs. Future work on this model will include member-check interviews with the participants to further investigate their pedagogical intentions when using examples in proof presentations.

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. (2009). Mathematicians' perspectives on the teaching and learning of proof. In Hitt, F., Holton, D., & Thompson, P. (Eds.), *Research in Collegiate Mathematics Education VII* (63-91). Providence, RI: American Mathematical Society.

Durbin, J. (2009). *Modern algebra: An introduction.* Sixth Edition. Hoboken, NJ: John Wiley & Sons.

Fukawa-Connelly, T. (2010). Modeling mathematical behaviors: Making sense of traditional teachers of advanced mathematics courses pedagogical moves. *Proceedings of the 13th Annual Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC.

Fukawa-Connelly, T., Newton, C., & Shrey, M. (in press). Analyzing the teaching of advanced mathematics courses via the enacted example space. To appear in *Educational Studies in Mathematics.*

Glaser, B. & Strauss, A. (1967). *The Discovery of grounded theory: Strategies for qualitative research.* New Brunswick, NJ: Aldine Transaction.

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, 185-212.

Harel, G., & Fuller, E. (2009). Contributions toward perspectives on learning and teaching proof. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 355-370). New York, NY: Routledge.

Harel, G., & Sowder, L (2007). Toward a comprehensive perspective on proof, In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, Reston, VA: National Council of Teachers of Mathematics.

Harel, G., & Sowder. L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in Collegiate Mathematics III* (pp. 234–282). Providence, RI: AMS.

Harel, G. & Sowder, L. (2009). College instructors views of students vis-a-vis proof. In M. Blanton, D. Stylianou, & E. Knuth (Eds.) *Teaching and Learning Proof Across the Grades: A K-16 Perspective*. Routledge/Taylor & Francis.

Hemmi, K. (2010). Three styles characterizing mathematicians' pedagogical perspectives on proof. *Educational Studies in Mathematics, 75*, 271-291.

Inglis, M., Mejia-Ramos, J., & Simpson, A. (2007). Modeling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics 66,* 3-21.

Lai, Y., Weber, K., & Mejia-Ramos, J. P. (in press). Mathematicians' perspectives on features of a good pedagogical proof. To appear in *Cognition and Instruction*.

Latakos, I. (1976). *Proofs and Refutations.* Cambridge: Cambridge University Press.

Lay, S. (2005). *Analysis with an introduction to proof.* Fourth Edition. Upper Saddle River, NJ: Pearson Prentice Hall.

Mason, J. & Watson, A. (2001). Getting students to create boundary examples. *Teaching and Learning Undergraduate Mathematics subgroup of the Mathematical Association Newsletter, 11*, 9-11.
Mason, J. & Watson, A. (2005). *Mathematics as a constructive activity: Learners generating examples.* Lawrence Erlbaum Associates: Mahwah, NJ.
Mejia-Ramos, J., 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.
Mejia-Ramos, J. & 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, 88-93, Taipei, Taiwan.
Michener, E. R. (1978). Understanding understanding mathematics. *Cognitive Science, 2*, 361–383.
Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education.* Portland, OR.
Polya, G. (1973). *How to solve it.* New Jersey: Princeton University Press.
Rosen, K. (2011). *Elementary number theory and its applications.* Sixth Edition. Boston, MA: Pearson.
Rowland, T. (2002). Generic proofs in number theory. In S.R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp.157-183). Westport, CT: Ablex Publishing.
Rowland, T. and Bills, L. (1999). Examples, generalization and proof. In Brown, L. (ed.) *Making Meanings in Mathematics: Refereed Proceedings of the British Society for Research into Learning Mathematics*. York, QED Publications, 103-116.
Speer, N. (2008). Connecting beliefs and practices: A fine-grained analysis of a college mathematics teacher's collection of beliefs and their relationship to his instructional practices. *Cognition and Instruction, 26*, 218-267.
Speer, N., Smith, J., Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior, 29*, 99-114.
Venema, G. (2002). *Foundations of Geometry.* Upper Saddle River, NJ: Pearson Prentice Hall.
Watson, A. & Mason, J. (2005). *Mathematics as a Constructive Activity: Learners Generating Examples.* Mahwah: Lawrence Erlbaum Associates.
Watson, A. & Shipman, S. (2008). Using learner generated examples to introduce new concepts. *Educational Studies in Mathematics, 69,* 97-109.
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*, 115-133.
Weber, K. (2011). Mathematicians' perspectives on their pedagogical practices with respect to proof. *Journal of Mathematical Education in Science and Technology,* DOI:10.1080/0020739X.2011.622803.
Yopp, D. (2011). How some research mathematicians and statisticians use proof in undergraduate mathematics. *Journal of Mathematical Behavior, 30*(2), 115-130.
Zodik, I. & Zaslavsky, O. (2008). Characteristics of teachers' choice of examples in and for the mathematics classroom. *Educational Studies in Mathematics, 69*, 165-182.

DISCURSIVE APPROACH TO STUDENTS' THINKING ABOUT THE DERIVATIVE

Jungeun Park
University of Central Arkansas

This study explores features of university calculus students' discourses on the derivative using a communicational approach to cognition. The data was collected from a survey and interviews in three calculus classes at a public Midwestern university. During the interview, 12 students explained their solution processes on the survey problems. The analysis of interviews focuses on students' descriptions about the derivative and the relationships between a function, the derivative function, and the derivative at a point. The results show that their descriptions were closely related to how they think about the derivative as a number and as a function. A common description of the derivative as a tangent line, which is a point-specific object but also a function defined on an interval, was identified. This description was closely related to their use of the word, "derivative" for both "the derivative function" and "the derivative at a point."

Key words: Communicational Approach to Cognition, Calculus, Derivative

Introduction

Research in collegiate mathematics education has been growing over the past few years, especially about calculus learning (e.g., Carlson, Oehrtman, & Thompson, 2008; Speer, Smith, & Horvath, 2010). Among calculus concepts, the derivative is known as a difficult concept because its definition contains various other concepts–ratio, limit, and function–and the derivative can be represented in multiple ways (e.g., Thompson, 1994; Zandieh, 2000). Related to previous studies, this study explores how students described the derivative and used the descriptions in tasked-based interview settings focusing on their use of the word, *derivative*. Unlike some languages (e.g., Korean or Japanese), *derivative* is colloquially used for both *the derivative function* and *the derivative at a point* in English. These two concepts are related, but mathematically different; the former is a function, and the latter is a number. This observation suggested an ambiguity about what *derivative* refers to and the possibility for miscommunication between speakers using the word, and provided a motivation for this study that addresses the following questions:

1. How do students describe the derivative at a point and the derivative of a function?
2. How do students use their concept of the derivative in problem-solving situations?

Here, problem-solving situations refer to task-based interview settings. Investigating students' thinking through their discourses can add new understanding to the current literature about the role that mathematical language plays in students' learning. There has been research about how word use is related to children's thinking about early mathematical concepts (e.g., Fuson & Kwon, 1992; Sfard, 2008), but few studies have been done in advanced concepts. An explanation about the use of the key words and visual representations may extend our understanding of the role that language plays in students' learning of an advanced concept, the derivative, and guide instructors' discourses about the derivative in class.

Theoretical Background

This study addresses students' thinking about the derivative at a point and the derivative of a function based on the mathematical relationship between function at a point and function on an interval. This section reviews existing literature reporting students' thinking about function and the derivative and addresses the characteristics of the mathematical discourses, which provided a theoretical lens for the analysis.

Function at a Point and Function on an Interval

There have been a rich body of research on how students understand function, which also have provided several ways to conceptualize the function. Especially, the studies, which address developmental stages of understanding functions, have made a clear distinction between the function at a point and function on an interval (e.g., Dubinsky& McDonald, 2001; Monk, 1994; Sfard, 1992; Breidenbach, Dubinsky, Hawks & Nichols, 1992). Most of these studies describe a first stage of understanding functions as being able to generate an output value of a function when an input value is given. For example, Monk (1994) called this view of function as "*pointwise understanding"* and described as making sense of a function that is used "as if it were only a table regarding particular" input values "as corresponding to particular" output values" (p. 21). A person in this stage would think of function as a value for a given input at a time. Sfard (1992) called this stage, "interiorzation," and Dubinsky and his colleagues (2001) called it "action." The next stage of understanding a function is described as being able to see dynamics of a function, i.e. all values at once. Monk called this stage "*across-time understanding,*" and described it as being able to explain how the changes in output variables and the change in input variables are related. Sfard (1992) called this stage "condensation," and Dubinsky and colleagues called it "process." These researchers also described later stages of understanding functions, but these are beyond the scope of Calculus I as Sfard (1992) mentioned that condensed concept of a function seems to be "sufficient…for…differentiation and integration" (p. 69).

Derivative at a Point and Derivative as a Function

Existing studies about students' thinking about the derivative can be divided in terms of the two types of understanding of functions described above. The studies about the derivative as a point-specific value include students' thinking about the limit of the difference quotient and the tangent lines to a curve. The results of these studies have shown that students' misconceptions about the limit (e.g., 0.999999…never reaches 1) (Tall & Vinner, 1981) are closely related to their thinking of the local linearity (Hahkioniemi, 2005) and the tangent line (e.g., the secant lines never reach the tangent line) (Tall, 1986; Orton, 1983). The studies about the derivative as a function mainly address co-variation. These studies have pointed out the importance of what is varying in a function. For example, Carlson, Oehrtman, and Thompson (2008) discussed that the rate of change of the volume of a sphere in terms of its radius is its surface area, but the rate of change of the volume of a cube in terms of its side is not the surface area. Thompson (1994) connected the concept of rate of change to students' thinking of the Mean Value Theorem.

However, few studies have been done about the relation between those two types of understanding of the derivative. Monk (1994) addressed these two types based on students' written answers on four survey problems including one derivative problem, but did not give detailed information about whether and how the students make a connection between these two concepts. This study expands these existing studies about students' thinking about the derivative focusing on how students describe and use the relationship between a function and its derivative.

Communicational Approach to Cognition

To explore students' discourse on the derivative, this study used the communicational approach to cognition (Sfard, 2008). This approach views thinking as an "individualized version of interpersonal communication" and mathematics as a discourse that is characterized by the four features: word use, visual mediator, endorsed narratives and routines. This study focused on the first three features. A *word* in mathematical discourse, which signifies mathematical objects, can be used differently in a different context (Sfard, 2008; Sfard & Lavie, 2005). For example, one word, "derivative" is used as the derivative at a point and the derivative of a function. When the word is used for the slope of a tangent line at a point, it refers to "the derivative at a point." When it is used in the context of differentiation rules, it mostly refers to "the derivative function" (e.g., Stewart, 2010). The words describe the concept of the derivative such as "slope" and "rate of change," and the quantifiers such as "a," "any," and "every," are also important to explore. *Visual mediators* refer to any non-verbal, visual objects used as a means of communication such as writing, drawing, and gestures. For example, usually the derivative at a point is denoted as $y = f'(a)$, and the derivative function is denoted as $y = f'(x)$ (e.g., Stewart, 2010). Here, "a" is used as a number, and "x" is used as a variable. The derivative at a point can be mediated by the graph of tangent line and the derivative function by its graph. This study focused on algebraic, graphical and symbolic notations that students used. *Narratives* are utterances that speakers can endorse as true or reject as false, and *endorsed narratives* refer to ones believed as true by speakers (Sfard, 2008, p.134). Students' endorsed narratives are often different from what the professional mathematics community endorses as true. Routines refer to meta-rules that determine discursive patterns. Because the interviews, which were conducted for an hour, did not provide enough information about the discursive patterns, routine was not included in the analysis.

Design of Study

This study is part of a larger study consisting of classroom observation, student survey, and interviews with instructors and students. Three calculus classes at a large public university in Midwest were observed for six weeks for the derivative unit. At the end of the unit, a survey was administered to the students in the classrooms, and then interviews were conducted with instructors and students after the survey. The students were selected for interviews based on their survey responses. Sfard's (2008) framework was used to analyze instructors' and students' discourses. This paper reports students' responses on the survey and their discourses during interviews. This section addresses a) survey and scoring, b) recruiting and interviewing students, and c) analyzing data. Results from the instructor interviews are reported in Park (2011).

The survey consisted of questions about students’ mathematics background and mathematical items involving a function, the derivative function, and the derivative at a point (See Appendix). Most items came from the Calculus Concept Inventory (Epstein, 2006), which included item reliability. Other items were reviewed by three mathematics professors. In the three classes, 88 of 99 enrolled students took the survey for 20 minutes in an exchange of 20 extra credit points out of 700 total. Two types of scores, raw and frequency, were calculated. Raw scores were based on correctness, and frequency scores were based on all students' responses in each class. For open-ended items, I coded students' responses into categories using the rubric I created. The maximum possible raw score was 23. For the frequency scores, I assigned 2 points for the most popular responses for an item (say n students select that response). If there was a response selected by more than $n/2$ students, I assigned 1 point for the response. If there were two (or more) most popular responses (say m students select each of those choices), I assigned 2 points for each response, and 1 point for a response that more than $m/2$ students selected. A student whose answers coincided with the most popular answers on all the problems received 32 points.

From each section, four students were invited for interviews based on their survey responses. The raw scores were used to find a heterogeneous group based on their survey performance. Frequency scores were used to find students whose answers were similar to the answers most commonly chosen by other students in the classroom. As shown in Table 1, most students interviewed from Instructor Alan's class had high raw scores (above 16 out of 23), most students interviewed from Instructor Ian's class had low raw scores (below 17), and there was a wide range of scores in interviewees from Instructor Tyler's class. Ten of the 12 students had studied the derivative in Advanced Placement Calculus in high school (Table 2).

Table 2. *Students Interviewed*

Instr-uctor	Name	Gender	Major	First Math Class Including Derivative	Raw Score	Frequency Score
Alan	Cole	M	Pre-med	Pre-calculus in HS	17	25
Alan	Zion	M	Chemical Engineering	Pre-calculus in HS	17	28
Alan	Bill	M	Engineering	Pre-calculus in HS	18	27
Alan	Joe	M	Civil Engineering	AP Calculus in HS	21	26
Tyler	Bob	M	Mathematics	Calculus I	17	20
Tyler	Liz	F	Med-Tech	Calculus in HS	11	22
Tyler	Zack	M	Computer Science	Pre-calculus in HS	15	26
Tyler	Neal	M	Computer Science	Calculus in HS	20	31
Ian	Sara	F	Biology	Calculus I	8	17
Ian	Mary	F	Genomics and Genetics	Pre-calculus in HS	13	21
Ian	Mona	F	Natural Science	Pre-calculus in HS	13	24
Ian	Clio	F	Astrophysics	Pre-calculus in HS	16	21

Note. In the table, AP, and HS refer to Advanced Placement and high school, respectively.

Task-based semi-structured interviews were conducted individually lasting for about an hour. During the interview, students were asked to answer warm-up questions about the derivative using their own words (Figure 1), and how they solved survey problems. Follow-up questions to their initial responses were focused on whether and how they used the relationships among a function, the derivative function, and the derivative at a point in their problem solving processes. Interviews were transcribed and coded with Transana (Woods & Fassnacht, 2007).

Q1. What is the derivative? Can you make a sentence with the word, "derivative"?
Q2. What is the derivative of a function?
Q3. What is the derivative at a point?
Q4. Is there any relationship between the last two terms?
Q5. Is a function related to the derivative of a function or derivative at a point?

Figure 1. Warm-up questions

Findings

This section addresses students' descriptions and uses of the derivative while answering the warm-up questions and explaining their solution processes. Their word use, visual mediators, and endorsed narratives were closely examined to explore their thinking about the derivative at a point, the derivative of a function, and their relation. This section only reports the cases that were identified at least in three different students' discourses, or three times in one student's discourse.

Word Use

While answering warm-up questions, most students (9 out of 12) explained the derivative using the phrases, "the slope" (Table 1). Other answers include "velocity" and "rules."

Table 1. *Students' Description of Derivative and Choice of f '(x) and f '(a)*

Name	Choice	The derivative of a function, *f '(x)*	The derivative at a point, *f '(a)*
		Tyler	
Roy	*f '(x)*	*"Extension or contraction of" f(x)*	Explanation was not provided
Liz	*f '(x)*	"How fast things change…Velocity over time"	"Velocity [or]…slope at a point"
Neal	*f '(a)*	"Graphical indication of every single point throughout a graph"	"Slope of a tangent line at a point" & "the direction the line is headed"
Zack	*f '(x)*	"Slope at all points" tells "what *f(x)* is doing"	"Slope at a point"
		Alan	
Bill	*f '(x)*	"You take it [*f(x)*] and derive the new equation…it describes…the slope"	Explanation was not provided
Zion	*f '(x)*	"Slope of the whole function at all the points"	"[Slope] at that single point."
Joe	*f '(a)*	"Slopes of the function at all points"	"Slope of a function at a point".
Cole	Both	"Rate of how fast something moves," & "Slope of the tangent line"	"it [these explanations] could be applied to both" *f '(x)* and *f '(a)*
		Ian	
Clio	*f '(x)*	"A way to…derive other equations from another using the rules"	"What the coordinates are at the given point…just plug it in"
Sara	Both	"The slope of the curve"	"Slope of the tangent line at a point"
Mary	*f '(x)*	"Short cuts" or "slope at all points"	"the slope of the function at a point"
Mona	*f '(a)*	"Slope of the whole line, another…function"	"Instantaneous slope at the point"

Of 12 students, eight students first said, "the derivative" as "the slope," one student as "velocity," and one student as "a way of deriving equations." Seven identified their description of "the derivative" as "the derivative function," with phrases "at all the points," "throughout the graph," and "over time." Three identified it as "the derivative at a point" with "at a point," and "at that single point." Two students stated that the same explanation for both concepts (e.g., "slope of the curve," & "slope of the tangent line at a point").

Students' use of the derivative as "slope" was also identified while students used the concept of the derivative to explain their solution processes on the survey problems. In contrast to their answers to the warm up questions, while they solved the problems, they did not use "slope" and "derivative," synonymously. For example, not all the students, who used "the slope" to find the graph of the derivative of a function that was given in problem 4, used the same interpretation to find the graph of the original function when its derivative function graph was given as in problem 5. For example, Bill supported his correct choice a) based on how "the slope" changes in problem 4 by saying, "because it [the graph of *f(x)*] has decreasing slope… getting less and less positive until it reaches negative." He, however, did not use the word, "slope" or the concept of the slope when he justified his choice e) in problem 5. He instead, found a similar shape of the graph of $y = f'(x)$. When they interpreted the derivative at a point in the problem context, $C'(2)$ in problem 1, most students (8) explained it as the change between $C(2)$ and $C(3)$ rather than the rate of change (e.g., "the cost to make another mile of rope," "how much more cost would be added for that one more unit…if you were to go to 3 from 2"). Two students initially answered, "marginal cost at $q = 2$," and "slope at $q = 2$" also explained it as the change in $C(q)$ between the two consecutive q values when I asked explain further. The units for $C'(2)$ was also consistent; of the eight students, five said, "dollars," and two said, "dollars/mile" but could not explain it in relation to their answer, the change between two consecutive q values. Although the change of a function between two consecutive x value is used as an approximation of the derivative at a point, their word use and the units imply their lack of understanding of the derivative as a rate of change.

Another phrase frequently used was "the derivative" as "tangent line." As we saw in Table 1, "tangent line" was included in most students' description of the derivative in warm-up session. However, in the problem solving process, they mentioned the derivative as "the tangent line" instead of its "slope." For example, out of 5 students who mentioned that the derivative as "the tangent line," two students—Bob and Joe—consistently used this concept in problem 8.

Visual Mediator

While explaining their concept of the derivative in the warm-up session, five out of 12 students used graphical representations; three drew a tangent line to a curve (Figure 2).

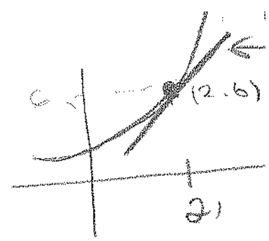


Figure 2. Students; Drawing for the Tangent line

Only two students drew two matching graphs of a function and a form derivative. Joe drew a curve for a function, and the tangent line at $x = 1$, and estimated its slope as 3, and plot (1, 3) on another x-y plane (Figure 3). Bill drew a graph of the derivative of $C(q)$ in problem 1.

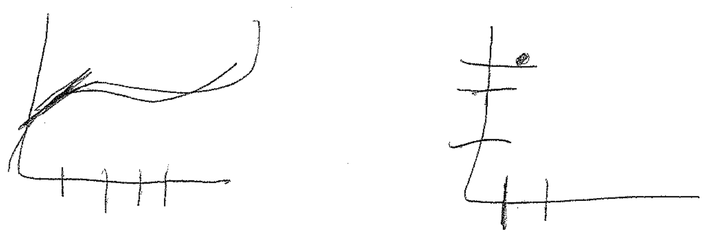
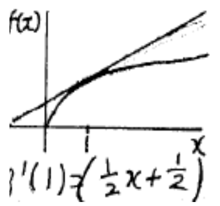


Figure 3. Student's Matching Graph of a Function and the Derivative at a Point

While solving problem 8, two students—Bob and Joe—drew a tangent line and said it is the derivative at a point. Joe also included the written notation $f'(1) = \frac{1}{2}x + \frac{1}{2}$ (Figure 3).



Joe : The tangent line at x=1, y = ½ x + ½, what this means is that $f\,'(1)$ equals 1/2x+1/2, so the function they gave you is particular to the point, x =1…If this only works if x=1, you need to have $f(1)$ somewhere…I didn't quite understand what these [choices] were trying to say. As far as manipulating this, I didn't see any $f(1)$…I can't really relate the function for a [tangent] line to the entire function of $f(x)$. It's only relevant at the point x = 1.

…

Joe: This is graph of f prime of x (adding a decreasing line to his graph). That tells you the value for the slope at that point. The slope would be 1, if you plug in 1 for ½ x + ½.

Interviewer: What you wrote here, '$f\,'(1)$ = ½ x + ½' would be $f\,'(x)$?

Joe: I don't think so. To say that $f\,'(x)$ equals something…the slopes [would] cover the entire domain…I don't think this [line]…has any other connection to the graph of $f(x)$ besides the slope at that one point.

Figure 4. Joe's Graphs of a Tangent line and $f\,'(x)$ and Explanation

As shown in Figure 4, Joe drew a decreasing line for $f\,'(x)$ and said, "this [line] tells you the value for the slope at that point." Later, he incorrectly found the slope of $f(x)$ at x = 1 by substituting x = 1 in the equation of the tangent line not in $f\,'(x)$, and gave another slope ½ from the equation, y = ½ x + ½. When I asked which one is correct, he chose latter but changed the answer by saying "the tangent line is a representative of the slope at this point…I guess that this whole thing [pointing to y = ½ x + ½] is the slope as opposed to just ½…it might be pretty wrong." This sentence indicates that his concept of the derivative as the slope includes the concept of the derivative as a function ("the whole thing"), and he considers $f\,'(1)$ as a point specific concept, but not as a number. In the same problem, two other students integrated the equation of the tangent line to find the equation of $f(x)$, which also suggests their inability to conceive of $f\,'(a)$ as a number.

The written notation of the derivative function given in Problem 9, "$f\,'(x) = ax^2 + b$" with words in the problem statement—"slope" and "tangent line"—seems to remind students of a linear function. Out of 12 students, 11 said "the slope is a" so it "should be positive" and "b can be any real number." After I asked them if "$x = 0$" matters or not in the solution process, seven of them were able to revise their solution correctly, but they also took several detours before they chose the correct answer.

Endorsed Narratives

Most frequently identified endorsed narrative, five students out of 12, was "the derivative increases/decreases if (or iff sometimes) a function increases/decreases." When I asked them to explain more, most of students corrected their statement, but two of them consistently used it to solve the problem. For example, a student, Bob, who consistently used this relationship between a function and the derivative function, made a connection to his use of "tangent line" as "the derivative" in word use and visual mediator in problem 8 (Figure 5).

e) None of these.

derivative function

8. **Consider the graph of f(x) below. The tangent line to this graph of $f(x)$ at $x = 1$ is given by $y = \frac{1}{2}x + \frac{1}{2}$. Which of the following statements is true and why?**

always increasing

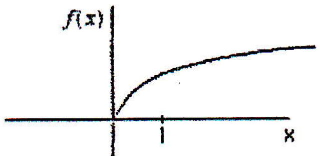


Figure 5. Bob's Solution in Problem 8.

Bob tried but failed to support his original choice "c) y ≥ ½x + ½" and change it to "e) None of these" by saying, because "y = ½x + ½ is the derivative of the function at that point," he "was not sure…to compare it to the whole function, *f(x)*." He consistently used "the equation of the tangent line" and "the derivative function at that point" synonymously to explain its behavior by saying, "it's always increasing, so is the function."

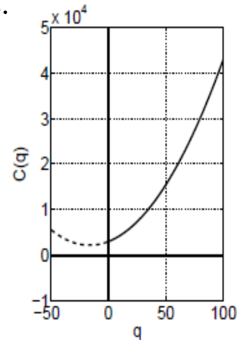
Discussion and Conclusion

This study contributes to the field of mathematics education by showing the importance of use of words and visual mediators in relation to students' thinking about the derivative. Existing research in this area has shown that students have various *misconceptions* of the derivative and some possible reasons (e.g., lack of understanding of the concept of limits and their procedural understanding of the rate of change). Some other studies related specific types of students' misconceptions (e.g., assuming resemblance in the graphs of $y = f'(x)$ and $f(x)$) to the limited contexts used in calculus books (e.g., increasing distance function whose velocity is also increasing). Research has also reported students' thinking about a function focusing on its co-varying nature (Monk, 1994; Thompson, 1994). This current study expands our understanding about the derivative by looking at the features of their discourses about the derivative. Mathematically, two terms, *the derivative of a function,* and *the derivative at a point* are consistent with *function* and *function at a point* because the derivative itself is a function. However, the results of this study showed students' lack of understanding this consistency, which was closely related to their use of the word, *derivative* and use of the visual mediator of the tangent line. Students showed a mixed concept of the derivative as a function defined on an interval, and a point-specific object simultaneously in graphical situations. This provides an explanation of their well-known misconception of the derivative as a tangent line (e.g., Zandieh, 1997). While describing or using this misconception, students used "derivative" without specifying the word as "the derivative at a point" and "the derivative function," which allowed them to change what the word referred to frequently even in one sentence. In their discourses, the word "derivative" was used not only as these two concepts, but also as "the tangent line" at a point. Also, students performed well on the items asking them to find the derivative at a point when the equation of the derivative of a function was given. However, their explanations on their solution process showed that they did not appreciate mathematical aspects behind the "plug-in" process or a sign of $f'(x)$ in relation to the behavior of $f(x)$ such as a) the derivative as a rate of change (the slope) describing the function behavior, b) the derivative at a point $f'(a)$ as a number, c) the derivative function $f'(x)$ as a function defined on an interval, and d) the relationship between $f'(a)$ and $f'(x)$: the former as a point-specific value of the latter. This lack of understanding was related to their incorrect endorsed narratives (e.g., if a function increases, the derivative increases) based on their concept of the derivative as the tangent line. These students' thinking about the derivative as a tangent line and the change in a function, might come from the confusion between what the derivative really is and its application. The tangent line is used as a linear approximation of a function (e.g., Newton's Method), and the derivative is also used to estimate change in function. These results suggest that calculus instructors should be careful about the use of the mathematical terms such as *function, the derivative, the derivative function,* and *the derivative at a point*, and visual mediators such as drawing and gesture for tangent lines. They need to be explicit about the mathematical aspects of these concepts especially when they introduce the concept of the derivative of a function and the derivative at a point, make a transition between these two concepts, and address what these two concepts represent in terms of the original function.

Appendix 1: Survey Questions

Please solve the following problems and show your work.

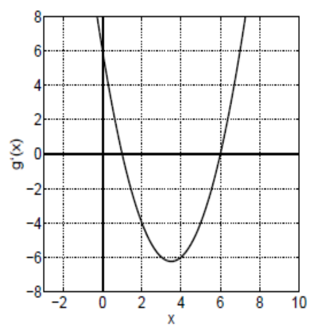
1. *C(q)* is the total cost (in dollars) required to set up a new rope factory and produce *q* miles of the rope. If the cost satisfies the equation $C(q)=3000+100q+3q^2$, and the graph is given as follows.



(a) Find the value of *C(2)*
(b) What are the units of 2 in (a)?
(c) What are the units of *C*(2)?
(d) What is the meaning of *C(2)* in the problem context?
(e) Find the value of *C(2)*.
(f) What are the units for 2 in (e)?
(g) What are the units of *C′(2)*?
(h) What is the meaning of *C′(2)* in the problem context?

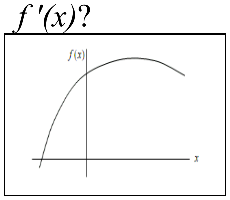
2. The derivative of a function *f*, is given as $f'(x) = x^2 - 7x + 6$. What is the value of $f'(2)$?

3. The graph of the derivative, *g′(x)* of function *g* is given as follows. What is the value of *g′(2)?*

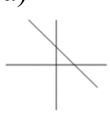


a) -4
b) -2
c) 0
d) 2
e) 4

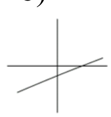
4. Below is the graph of a function *f(x)*, which choice a) to e) could be a graph of the derivative, *f ′(x)*?



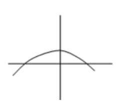
a)



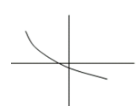
b)



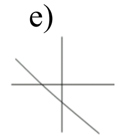
c)



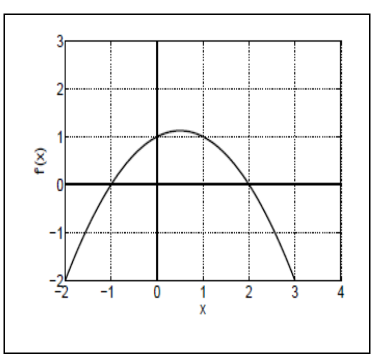
d)



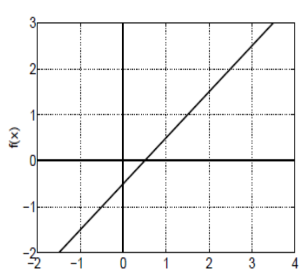
e)



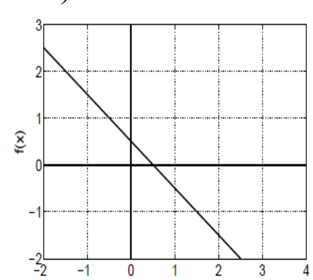
5. Below is the graph of the derivative *f ′(x)* of a function *f(x)*. Which choice a) to e) could be a graph of the function *f(x)*?



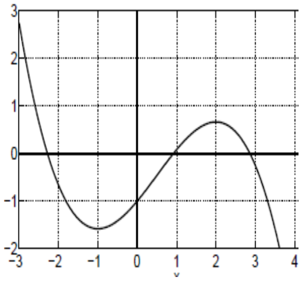
a)



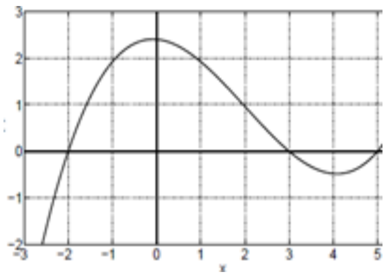
b)



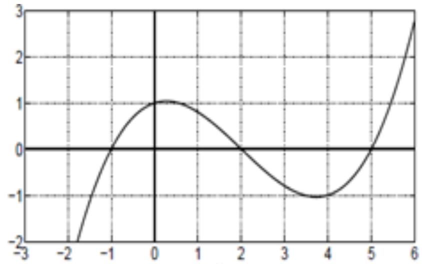
c)



d)



e)

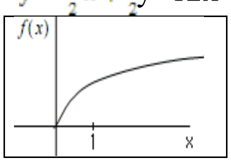


f) None of these

6. If a function is always positive, then what must be true about its derivative function?
   a) The derivative function is always positive.
   b) The derivative function is never negative.
   c) The derivative function is increasing.
   d) The derivative function is decreasing.
   e) You can't conclude anything about the derivative function.

7. The derivative of a function *f(x)* is negative on the interval *x*=2 to *x* =3. What is true for the function *f(x)*?
   a) The function *f(x)* is positive on this interval.
   b) The function *f(x)* is negative on this interval.
   c) The maximum value of the function *f(x)* over the interval occurs at *x*=2.
   d) The maximum value of the function *f(x)* over the interval occurs at *x*=3.
   e) We cannot tell any of the above.

8. Consider the graph below. The tangent line to this graph of *f(x)* at *x* = 1 is given by $y = \frac{1}{2}x + \frac{1}{2}$y=12x+12. Which of the following statements is true and why?



a) $\frac{1}{2}x + \frac{1}{2} = f(x)$ b) $\frac{1}{2}x + \frac{1}{2} \le f(x)$ c) $\frac{1}{2}x + \frac{1}{2} \ge f(x)$ d) $\frac{1}{2}x = \frac{1}{2}f(x)$

e) None of these

9. The derivative of a function, *f,* is $f'(x) = ax^2 + b$. What is required of the values of *a* and *b* so that the slope of the tangent line to the function *f* will be positive at *x* = 0.
   a) *a* and *b* must both be positive numbers.
   b) *a* must be positive, while *b* can be any real number.
   c) *a* can be any real number, while *b* must be positive.
   d) *a* and *b* can be any real numbers.
   e) None of these
   Why?

References

Bezuidenhout, J. (1998). First-year university students' understanding of rate of change. International Journal of Mathematics Education in Science and Technology 29(3) 389-399.

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., Oehrtman, M. C., & Thompson, P. W. (2008). Foundational Reasoning Ability that Promote Coherence in Students' Function Understanding. In M. P. Carlson & C. Rasmussen (Eds.), Making the connection: Research and practice in *undergraduate mathematics* (pp. 150-171). Washington, DC: Mathematical Association of America.

Dubinsky, E., & McDonald, M. A. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research in D. Holton, et al. (Eds.), *The Teaching and Learning of Mathematics at University Level: An ICMI Study* (pp. 273-280). Kluwer Academic Publishers, Dordrecht, Netherlands.

Epstein, J. (2006). "The Calculus Concept Inventory," *Proceedings of the National STEM Assessment Conferenc*e, Washington DC, October 19-21, 2006 p. 60 - 67

Ferrini-Mundy, J., & Graham, K. (1994). Research in calculus learning: Understanding of limits, derivatives, and integrals. Research Issues in Undergraduate Mathematics learning; Preliminary Analyses and Results. MAA notes 33, 31-45.

Fuson, K., & Kown, Y., (1992). Korean Children's Understanding Multidigit Addition and Subtraction, *Child development,* 63(2), 491-506.

Hauger, G. (1998). *High school and college students' knowledge of rate of change.* Unpublished doctorial dissertation, Michigan State University, Michigan.

Monk, S. (1994). Students' understanding of functions in calculus courses, *Humanistic Mathematics Network Journal,* 21-29.

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

Santos, A. Thomas, M. (2003). Representational ability and understanding of derivative. Proceedings of the 2003 Joint Meeting of PME and PMENA, 2, 325-332.

Sfard, A. (1992) 'Operational origins of mathematical objects and the quandary of reification-the case of function'. In Harel, G. and Dubinsky, E. (Eds), *The Concept of Function: Aspects of Epistemology and Pedagogy, MAA Notes 25* , pp.59-84, Washington:MAA

Sfard A. (2008). *Thinking as communication.* Cambridge University Press

Sfard, A. & Lavi, I. (2005). Why cannot children see as the same what grownups cannot see as different? – early numerical thinking revisited. *Cognition and Instruction*, *23(2)*, 237-309

Speer, N. M., Smith, J. P., Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, 19, 99-114

Stewart, J. (2010) *Calculus.* Mason, OH: 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(2) 151-169.

Tall, D. O. (1987). Constructing the concept image of a tangent. *Proceedings of Eleventh International Conference of Psychology of Mathematics Education,* (Vol. 3, pp. 69-75). Montreal, Canada.

Thompson, P. (1994). Images of rate and operational understanding of the fundamental theorem of calculus, *Educational studies in mathematics* 26 (2/3), 229-274.

Woods, D, & Fassnacht, C. (2007). Transana v2.20. http://www.transana.org. Madison, WI: The Board of Regents of the University of Wisconsin System.

Zandieh, M. (1997). *The evolution of student understanding of the concept of derivative.* Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

Zandieh, M. (2000) A theoretical Framework for analyzing students understanding of the concept of derivative. CBMS *Issues in Mathematics Education*, 8, 103-127.

WHAT'S THE BIG IDEA?: MATHEMATICIANS' AND UNDERGRADUATES' PROOF SUMMARIES

Aron Samkoff Rutgers University
Keith Weber Rutgers University
Pablo Mejia-Ramos Rutgers University

In this study, seven mathematicians and seven undergraduates were asked to read and summarize mathematical proofs that they read to investigate which ideas they consider to be important in a proof. Mathematicians' ideas consisted of a) the overarching goals of the proof, b) the ideas they found novel or unfamiliar, c) theorems or facts used in the proof, d) encapsulations of inferences as applications to general methods, e) diagrams, or f) cues for reconstructing the proof. Students did not mention the goals, important theorems or facts, or the methods as being important, but some focused on whether the proof was indirect or direct. We present a model for accounting for many of the different types of ideas found important by mathematicians.

Key words: Proof, key ideas, proof reading, proof summaries.

1. Introduction

1.1. The importance of reading and understanding proofs

Proof plays a central role in upper-level mathematics. In advanced mathematics classes, a great deal of time is spent presenting proofs of mathematical theorems to students. Despite this, there has been comparatively little research on how individuals read proofs. Of the work that has been done, most of it deals with how students validate proofs that they have read (e.g., Alcock & Weber, 2005; Selden & Selden, 2003). In particular, there has been very little research on how students understand the proofs they read (see, e.g., Mejia-Ramos, 2008).

1.2. What does it mean to understand a proof?

By comprehension of proofs, we mean that the reader engages with the text with the goal of *learning* from it, as opposed to evaluating it for correctness. Mejia-Ramos (2008) conducted a search of the literature and found that of the articles written on proof, the vast majority focused on tasks involving proof *construction*, with fewer on proof *presentation* and proof *reading*. Of those focusing on *reading*, only three dealt with proof comprehension. Conradie and Frith (2000) suggested assessing students' understanding of proofs through asking them to read proofs and answer related questions. Yang and Lin (2008) put forth a model for high school students' comprehension of geometry proofs that consisted of moving through various 'levels' of understanding. Based on these works and other suggestions from the literature, Mejia-Ramos et al (2012) put forth a general model for undergraduates' comprehension of mathematical proof consisting of seven orthogonal 'facets' of understanding, including one relating to the summarizing of the high-level ideas contained in the proof. Mejia-Ramos et al (2012) claimed that students' ability to identify or construct a summary of a proof constituted a measure of their understanding. We wish to investigate this facet of understanding further.

1.3. "The crux" of a proof in mathematics education

We are interested in how individuals construct summaries of proofs. As we discuss in the following section, summarizing includes identifying the *important ideas* of a proof. Several researchers have attempted to define "the crux" of a proof. Raman (2003) defined the "key idea" of a proof to be the mapping from an informal, private argument to a public proof. Hanna (1990) distinguished "proofs that explain" from proofs that merely prove by the fact that the former relied on a specific mathematical property. Rav (1999) states that the important ideas contained in proofs are not the theorems themselves, but the methods that mathematicians can use in other situations. Leron (1983) suggested that proofs be presented in a way that emphasizes their high-level ideas.

These reveal a multiple perspectives on what should constitute the important ideas of a proof. Furthermore, some of these suggestions are not fully operationalized. For example, some researchers have claimed it is difficult to distinguish a proof that explains from a proof that does not (Raman, 2003). Finally, these suggestions have not been empirically studied. One of our research goals is to gain empirical support for the existence of these types of ideas by studying what mathematicians find important when they read proofs.

2. Theoretical perspective

2.1. Reading as a constructive activity

There is a consensus amongst researchers in reading comprehension that reading is not a passive activity in which the reader absorbs the meaning contained in a text, but rather an active one in which the reader actively constructs meaning from the text by making inferences and relating ideas in the text to other parts of the text and his or her own background knowledge (see, for example, Dole et al, 1991). In this view, readers' background knowledge is as important as what is contained in the text, and comprehension requires active sense-making on the part of the reader (Duke & Pearson, 2002). One way in which to gain insight into how individuals understand from reading a text is to ask them to *summarize* the text. Although commonly used as an assessment tool in reading comprehension, this element of understanding is yet unexplored in mathematics education.

2.2. Summarization as a part of comprehension

Dole et al (1991) define summarizing a text to be an activity that requires the reader to identify the important ideas of a text that are then composed to form a new text that captures the meaning of the original text. Good readers often construct summaries of texts they have read (e.g., Duke & Pearson, 2002). Furthermore, there has been success in teaching students to summarize texts (e.g., Cunningham, 1982). Importantly, students who received instruction in summarizing not only became better at summarizing, but improved in other measures of comprehension. Thiede and Anderson (2003) gave empirical support to this argument by showing that participants who summarized a text more accurately gauged their understanding as compared with a control group. In sum, research suggests that summarizing is linked to understanding. Therefore, summarizing is an activity that allows a researcher to assess a reader's understanding of a text through what they find to be most important in the text.

2. 3. Proof reading as a constructive activity

Like reading any text, we contend that reading proofs also is a constructive activity that involves active engagement on the part of the reader. For example, Weber and Alcock (2005) argued that actively inferring and checking warrants is an important part of the proof reading process. Weber and Mejia-Ramos (2011) documented that mathematicians use examples and construct sub-proofs in order to comprehend proofs that they read, and that mathematicians' goals when comprehending a proof include understanding the proof at a high level, in addition to checking warrants. This provides a theoretical reason to believe that summarizing a proof would be related to understanding—summarizing may help readers gain insight into the high-level idea of the proof.

3. Methods

3.1 Participants

Seven mathematicians and seven undergraduates were invited to participate in the study. All participants were recruited from a large public university in northeastern United States, and all who were invited agreed to participate. Six of the seven mathematicians were tenured faculty, and one was an advanced graduate student who would soon successfully defend his doctoral dissertation, and was subsequently employed as a post-doctoral fellow at a different institution. All seven mathematicians were recruited because the first author knew them to be thoughtful and articulate when discussing mathematics.

The undergraduates were mathematics majors, and were recruited from a real analysis course at the university. All students had completed an introductory mathematical reasoning course in which students are typically taught elementary proof writing techniques, and basic set theory. They were paid a modest fee for their participation.

3.2. Materials

Three of the four proofs used in this study are presented in the Appendix. These proofs were designed to have the following features: These proofs were chosen because they were sufficiently long and complex so that they could be summarized, yet still would be accessible to an undergraduate audience.

3.3. Procedure

Participants met individually with the first author of the paper for a one-hour task-based interview. Participants were told to think aloud while reading each proof, and were then prompted to write a summary of the proof for themselves. Researchers in reading comprehension have noted that asking individuals to provide personal summaries provided insight into the ideas of the proof that they found important while asking individuals to write summaries for a different audience tended to lead them to emphasize ideas that they believed the researcher or evaluator would find important (Dole et al, 1991; Hidi & Anderson, 1986) Given these findings, the interviewer emphasized to participants that their task was to produce personal summaries.

After producing a summary of the proof, participants were asked to describe their summaries; they were also asked what they thought were the most important ideas in the proof. They were also asked to provide further detail or clarification for any behaviors that the researcher found interesting, confusing, or important. After these questions were answered, participants were presented with another proof that they were asked to summaries. This process continued until participants summarized all four proofs, or until 45 minutes had elapsed.

Next, participants engaged in an semi-structured interview in which they were asked the following questions:

-How would you define an important idea in a proof? Are there different types of important ideas that might appear in a proof?

-Is summarizing a proof something that you do in your mathematical work/studies?

-For what purposes do you summarize a proof? If so, how would your summaries change depending on the purpose?

-Are there different audiences for whom you would summarize a proof? If so, how would your summaries change depending upon the audience?

3.4 Analysis

The primary data source for the analysis was participants' descriptions of their summaries and what parts of the proof that they believed were important.

Analysis consisted of two phases. In the first phase, all interview data were transcribed and coded for common themes using a grounded theory approach (Corbin & Strauss, 2008). Preliminary definitions for classes of important ideas in proofs were formed using the mathematicians' interview data. In the second phase, mathematicians' written summaries were examined to capture specific instances of the themes discussed, and to refine the categories that were already created. Students' transcripts and summaries were then examined using the codes created from the mathematicians' data. The result of this analysis was a table of the different types of important ideas mentioned by students and mathematicians with a list of the participants who discussed each type of idea.

4. Results

The types of ideas participants viewed as important are presented in Table 1.

Type of important idea	**Mathematicians who mentioned this idea was important**	**Students who mentioned this idea was important**
Goals	M1, M2, M3, M4, M5, M6, M7	U1, U7
Novel/unfamiliar ideas	M1, M2, M3, M4, M7	U1, U2, U4, U5, U6, U7
Important Theorem/Fact	M1, M2, M5, M6	U1
Encapsulation of inferences	M2, M4, M6	---
Diagram/Graph	M2, M6, M7	U2
Proof techniques	---	U2, U4, U5
Cues for reconstruction	M1, M2, M4, M7	U5, U6, U7

Table 1: A summary of the types of ideas participants viewed as important in a proof.

In what follows, we describe each of the categories above and give instances of each from participants interview data and written summaries.

4.1. Goals

An excerpt was coded as "goals" whenever a participant mentioned the broad steps taken by the proof to be important. All seven mathematicians mentioned this type of idea as being important in a proof. An excerpt from M5 illustrates this:

> M5: So the important idea is, what are we going to do next? What's the goal, or what is the short-term goal in getting toward the final assertion...How are we going to successively reduce—you know, make some progress from our starting point to the finishing point? So there will be certain observations that can be made, sometimes a problem can be reduced from one thing to another, or you say it's enough to know something, for example, in the second problem, if we can see that the derivative is always positive, then we know there is at most one solution.

Another mathematician, M4, mentioned attending to the sufficient conditions set up in Proof 3 to establish the claim:

> M4: Yeah, you're trying to show under certain hypotheses that something is true, and typically the proof, what happens is that you wanted to show that this conclusion, I'll call it C, is true, and you need to say, well what are the properties that would imply C? Sometimes it's just a definition, you know, the property C means a certain thing. But sometimes you have some general sufficient condition for a property C…*So right, any time that you kind of switch your focus to a new goal and that you show that this goal is sufficient for that, those are really the important or key ideas* (Italics added for emphasis.)

As an example of an instance in which a mathematician wrote down the goals of the proof in his or her summary, M3 highlighted the goals taken in Proof 1 and 2 in his written summary, as shown in Figure 1 and Figure 2. In each of these summaries, M3 appears to reduce each proof to two broad steps. The fact that these steps are numbered and that they are all M3 wrote down as a summary suggests that he views this summary as a skeleton of the overall structure of the proofs. In section 5, we present a model to further explain this occurrence.

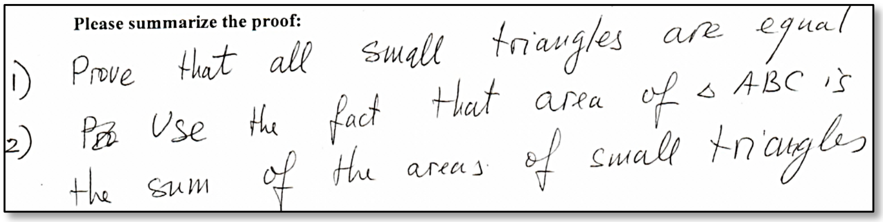

Please summarize the proof:
1) Prove that all small triangles are equal
2) ~~P~~ Use the fact that area of $\triangle ABC$ is the sum of the areas of small triangles

Figure 1: M3 highlights the goals in Proof 1.

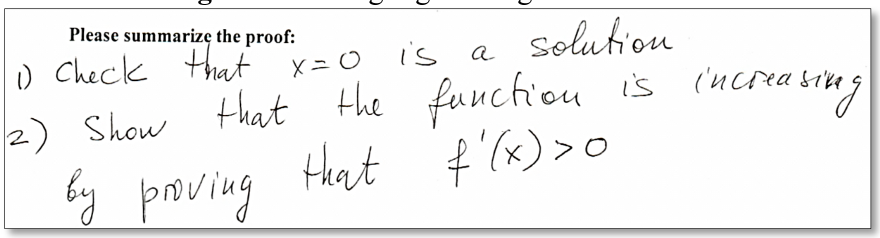

Please summarize the proof:
1) Check that $x = 0$ is a solution
2) Show that the function is increasing by proving that $f'(x) > 0$

Figure 2: M3 highlights the goals in Proof 2.

By contrast with the mathematicians, only two students mentioned this type of idea to be important in a proof.

4.2 Novelty

An excerpt was coded under this category whenever a participant mentioned including ideas that were novel, surprising, or unfamiliar to him or her. Five mathematicians viewed this type of idea as important in a proof. A representative comment was made by M7:

> I: How do you define an important idea of a proof?
>
> M7: It's possibly something that I wouldn't have thought of on my own. Not immediately obvious to do. That's what I would definitely have to write down in a summary…One of the most important things are the—perhaps the most surprising ideas. Here, these were all basic stuff that we could do in an intro to proof course, nothing was really surprising there. When I'm reading a research paper or something, and I have to mark the more surprising ideas that I would not have come up with on my own. Or they reference theorems I had never heard of before.

Smaller scale examples of novel ideas can be found on examining M4's experience reading Proof 3. As presented in section 4.1, M4, whose field of research is combinatorics, identifies a strategy employed in Proof 3. This proof is taken from the area of number theory, and shows that two numbers are perfect squares by showing that they are coprime and that their product is a perfect square. M4 suggests that he would not necessarily know to use this approach before reading the proof:

> M4: So I think that....that idea, that particular strategy is one—and it's important to me because I don't usually do number theory proofs, and so while this particular strategy is clear once it's stated, it's not necessarily a strategy that would jump out at me. So yeah that's an important idea.

Later on in the proof, the two numbers are shown to be coprime by assuming k divides both numbers and showing that $k = 1$. M4 comments on this approach as well:

> M4: But this idea of saying, oh, but then you can also use the fact that if k divides this and k divides that, then k^2 divides this squared, and k^2 divides that squared, and you can apply that. So that is kind of clear, except it's not something that I automatically think of as something to try.

In both instances, M4 acknowledges that he would not have thought to use the particular strategies employed in the proof, and consequently, M4 considered these ideas to be important.

While mathematicians tended to emphasize ideas that were new to them, M1, M2, M3, M4, and M7 all mentioned *omitting* ideas from their summary that they found to be routine, automatic, or trivial to them. M1 discusses the distinction between summarizing proofs containing familiar and unfamiliar ideas in the context of his own practice as a mathematician:

> M1: It depends on how comfortable I am with the material. So I mean if I'm lecturing on something that I know well, and I'm sure I know how to do, a summary will just consist of maybe the statements of things I want to prove. Or just a short indication of the proof. If I'm reading somebody else's paper and trying to convince myself that I understand the argument, then I'll tend to write down stuff that's very detailed, sort of checking every point, until I suddenly realize OK, I understand what they're doing, and this is how it goes, and everything else is OK, and then I may just put 'dot dot dot'.

In the above excerpt, M1 asserted that his summary of a proof would depend on his familiarity with the material, suggesting that there might not be agreement in mathematicians' personal summaries of a proof, as what is familiar to one mathematician might be surprising to another.

4.3 Important theorem or fact

An excerpt was coded in this category whenever a mathematician stated that a particular theorem or fact was crucial to a given proof. Four mathematicians made comments that were coded under this category. Of these, M2, and M5 made reference to the use of Rolle's Theorem in Proof 2; M1 and M5 referred to the fact that the Pythagorean relation in Proof 3 can be factored, while M6 mentioned theorems or facts as being important without referring to a specific instance in the proofs she read. For example, M2 states:

> M2: The important idea is that I'm going to use Rolle's Theorem, which is essential in transporting information between a function, its derivative, and their roots. Uh, so something's important if…it encapsulates a whole bunch of of examples. Very, very, very conveniently. OK, I think I could prove that Rolle's Theorem example without citing Rolle's Theorem. But it would be incredibly awkward and tedious. It would just—take me 35 minutes to come up with something…and it would be unpleasant and awkward and I just—I like the idea of having a shiny box labeled "Rolle's Theorem"

In this excerpt, M2 claims that the application of Rolle's Theorem to the function in Proof 2 is important because Rolle's Theorem is a powerful tool which makes the argument relatively easy. With regard to Proof 3, M1 said the following:

> M1: The real key thing here is that c^2-a^2 equals b^2. And that you can factor c^2-a^2 as c+a times c-a. So I started out my summary with that.

However, only one student mentioned theorems or facts as being particularly important in the proofs they read.

4.4 Encapsulation of inferences

An idea was coded as an encapsulation of inferences whenever a participant mentioned viewing a series of calculations or inferences as the application of a single technique. Three mathematicians mentioned this type of idea as being important in a proof. For example, M2 identified a series of calculations in Proof 2 as being an instance of "completing the square", and compared this to other settings in which he has used this technique:

> M2: In proving the Cauchy-Schwarz inequality, which is what you prove in, let's see, I'll tell you the settings I've proved it in. Complex analysis. Real analysis. Functional analysis. Various undergraduate courses. In all of these settings, they're different settings for the Cauchy-Schwarz inequality. The proof is, complete the square…So computational schemes like complete the square can be important.

The method of completing the square specifically played a role in M2's summary of Proof 2, shown in Figure 3. In this summary, M2 highlights the method of completing the square in addition to the goals of the proof. He then verbally summarizes the proof: "Rolle's Theorem applied to x^3 + 5x = 3x^2 + sin(x). Zero is a root. x^3 +5x-3x^2-sin(x) has a derivative which is never zero. Complete the square. That's what I'd write."

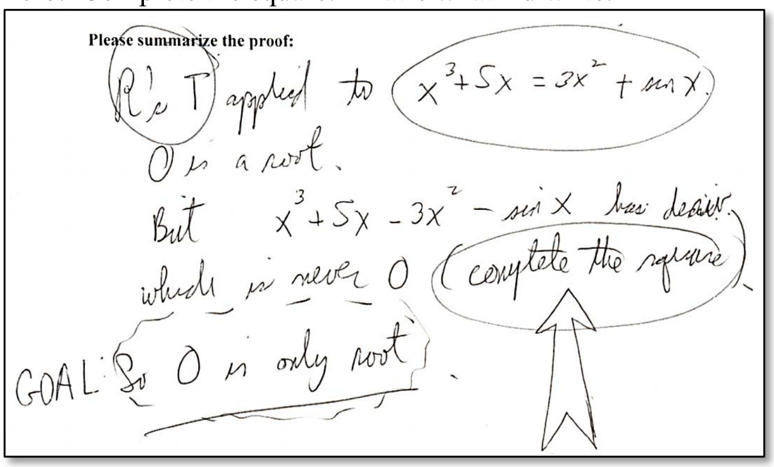


Figure 3: M2's summary of Proof 2. Note the emphasizing of the important ideas of the proof.

4.5 Diagram/graph

An excerpt was coded under this category whenever a mathematician mentioned that a diagram represented an important idea in a proof. This occurred with three of the mathematicians. The following excerpt from M2, in which he refers to his summary of Proof 1 (see Figure 4), illustrates the idea:

> M2: My summary…The picture above, and area DFE equals ¼ area ACB. My summary would have been exactly that. The first proof you gave me. The diagram, to me, is the—*yeah, the diagram's the important idea of the proof*! To me personally, the reasoning is a very pale reflection of the geometric impact of the diagram. The reasoning sort of writes down almost painfully what the diagram reveals more or less instantaneously.

Although M2 does not elaborate on how or what the diagram "reveals instantaneously", there has been ample research on diagram usage in mathematics to explain what he might mean . For example, Raman and Weber (2006) suggested that a diagram can reveal properties that would be otherwise be non-obvious. This would account for M2's quote, since the diagram would provide a skeleton of statements which could then be successively justified by deductive arguments. We discuss this in more detail in section 5.

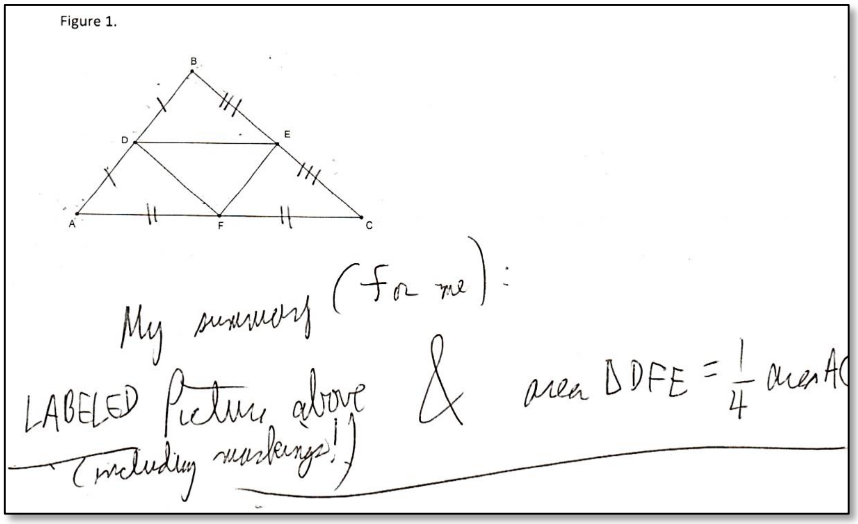


Figure 4: M2's summary of Proof 1.[1]

4.6 Proof techniques

Three undergraduates mentioned the proof technique (e.g., proof by contradiction) employed in the proof to be an important idea. U4 illustrates this idea in the following excerpt:

> U4: Yeah, or is it directly just a usual proof, because using certain theorems, like this theorem, that theorem, or something, they really just conjecture this happens, that happens. Like this implies that, this implies this, that kind of thing. So you're just going to sort of use definitions to apply different things to get to a solution, fine. Or are you going to say, OK, well we're going to do a contradiction. But though, then, the beginning, I think is helpful. Because at least you know what the framework of the whole proof is going to be.

By contrast, no mathematicians mentioned the indirect proof employed in Proof 2 or 4 as being important.

4.7 Cues for reconstruction

An excerpt was coded under this category whenever a participant mentioned a particular piece of information as being important either for reminding themselves of the approach taken in the proof, or for saving themselves time or effort in reconstructing parts of the proof. Three mathematicians mentioned including ideas for this reason. M4 illustrates this way of thinking in the following quote:

> M4: So now I'm realizing that it's not so much—what I'm writing down is not so much the important ideas, it's the ideas that...sort of don't come to my head automatically. So if there's something that when presented with this I have to play around a little bit before I would see it, then I would write the note... That makes it an important idea, but… I haven't really thought about that. When I'm writing these notes, it's really focusing on just the points in the proof where there's a choice to be made, or some search among different strategies for the particular step you're about to take. And the strategy that's used is not the first strategy that I think about.

In this way, M4 appears to view these types of ideas to important simply for the sake of saving time by avoiding re-inventing the steps of the proof. M1 mentioned a similar idea

---

[1] M2's writing was cut off as he started writing on another page. The line on the bottom right reads: □□□□Δ□□□ *=14□□□□(□□□)* .

after writing the summary for Proof 2, concerning the existence of a solution of a transcendental equation:

> M1: *So the way you could waste a lot of time on this problem is to try and find a solution*. So if the sin(x) hadn't been there, then this would have been a cubic expression with zero constant term, so you could say, ok you can factor out an x, and you get a quadratic, and the roots are such and such. And zero is the only real root. So you do it all explicitly. And it's always tempting to try and do that. (Italics added for emphasis.)

When asked about the important ideas contained in Proof 2, he expands on this idea:

> M1: *Well the key is sort of a negative thing that you don't try and solve the equation*, but you write it as a function equal to zero and want to show for some reason, there's only one root and well if you can show that it's strictly increasing that'll do it. And so you take the derivative…*So the key idea there actually is negative—it's that you don't try and write down a root.* I mean you check that zero is a root, just to make sure whoever set up the problem isn't lying to you, but you don't try and explicitly write down another root. (Italics added for emphasis.)

Similarly to M4, M1 states that he is considering the time and effort that would be wasted in trying to re-invent the method of the proof from scratch. We argue that part of the role of the summary for these mathematicians was to provide a scaffold for reconstructing the proof on their own.

5. Discussion

5.1. What mathematicians and students valued in proofs

Table 1 provides a typology of the different types of ideas the mathematicians in this study found important. That the mathematicians emphasized what was novel to them and omit what is routine suggest an important difference between summarizing a proof versus a general (non-mathematical) text. In non-mathematical texts, there can be information that is redundant or tangential to the main ideas of the text that can be omitted without changing the central meaning of the text or the validity of an argument. Proofs seldom contain such information. Thus, in deleting information in forming a summary, one cannot use the criteria of redundancy or extraneous, but the more subjective notion of familiarity with ideas. This would imply that this aspect of summarizing a mathematical proof is necessarily subjective. We have also identified differences between what the students and mathematicians in our study. First, students found the proof technique to be important, but mathematicians did not. There are at least two ways to account for this difference. Perhaps students' unfamiliarity with proof by contradiction led them to place importance on this technique. Indeed, this seems to be the case with U5, who claimed that he explicitly attends to the proof technique while reading the proof:

> U5: Uh taking [real analysis] now, one of the things that I wonder about is how you're supposed to know when to use a direct proof and when to use a proof by contradiction or something like that, so any time I see one that uses a proof by contradiction, I kind of try to figure out, all right, why –how do you know beforehand that that's the way to do it?

It is plausible that U5's inexperience with indirect proof that leads him to pay such explicit attention to it. A mathematician would be very familiar with indirect proof, and would therefore not attend so explicitly to its use. Another explanation is that some students might view proofs by contradiction as being somehow essentially different from direct proofs, whereas mathematicians generally do not consider the proof technique used as being essential part of a proof.

Second, the students did not find the goals or encapsulations of inferences to be important in proofs. One explanation for this is that the students generally did not consider the high-level ideas when reading these proofs. That is, they may have been focusing on how each step followed from the previous, without doing what Weber and Mejia-Ramos (2011) refer to as "zooming out". This is consistent with Selden and Selden's (2003) observation that students focus locally on calculations when validating a proof but ignore the proof's overarching structure.

Finally, the students did not generally mention important theorems in their interviews. A possible reason for this is that students lack the experience that mathematicians do in applying ideas in a number of different settings, and thus fail to appreciate such theorems' power and applicability.

5.2. A model for how mathematicians summarize proofs

In this section, we present a model for how mathematicians summarize proofs which accounts for the descriptions given by the mathematicians in this study. This model is based on the model for a written proof found in Rav (1999) and Arzarello (2007). In this model, a proof is viewed as a sequence of statements A_0 through A_n, with A_0 representing the hypothesis or hypotheses, and A_n representing the conclusion and each arrow representing the warrant or argument that allows each A_{i+1} to be deduced from A_i. We note that this model is of course a simplification of reality—mathematicians may often read proofs that have much more complex structure . However, we believe this is a useful simplification in order to understand the behavior of mathematicians reading relatively simple proofs such as the ones used in this study.

Given a proof in this linear format, if a hypothetical mathematician were asked to summarize this proof, she would likely identify the hypotheses and conclusion statements of the theorem statement (represented by the top- and bottom-most nodes in our model). We propose that she then lays out the structure of the argument by identifying key statements achieved in the proof. She can do this in two ways: 1) by identifying the goals achieved in the proof, or 2) by looking at a diagram which makes certain statements of the proof immediate. These help her identify "landmarks" or "guideposts" (represented by the intermediary nodes in our model) that guide her through the proof. However, these guideposts may not be enough for her to reconstruct the proof on her own. She might also need to remind herself *how to proceed from one guidepost to the next*. She might do this for a given part of the proof by 1) encapsulating inferences as the application of a general technique, 2) identifying the novel ideas or techniques, 3) identifying an important theorem or fact, or 4) providing a cue for reconstructing a part of the proof.

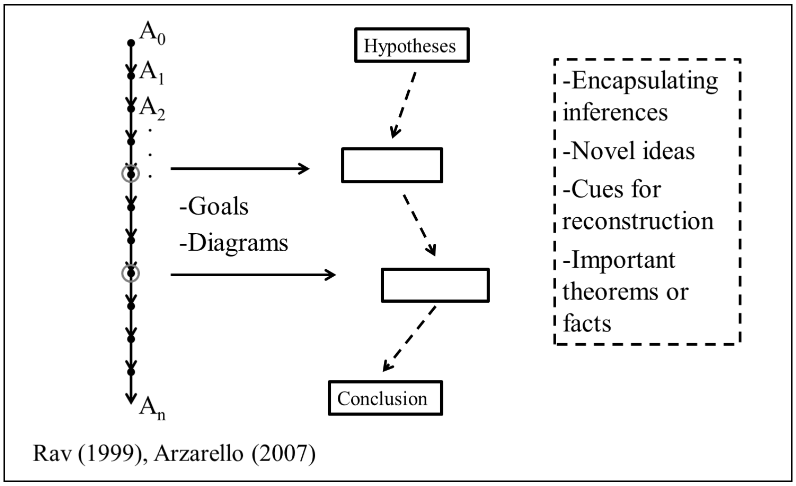


Figure 5: A model for how mathematicians summarize proofs.

5.3. A more nuanced view of "the crux" of a proof

The results presented in section 4 suggest that the notion of "the key idea" or "the crux" of a proof is somewhat complex. First, there appears to be no universal type of important idea that occurred in all of the proofs. Rather, mathematicians had diverse views on what could constitute an important facet of a proof and these might vary by the proofs that they read. Some mathematics educators have attempted to define "the key idea" of a proof (e.g., Raman, 2003). As we argue below, we believe to have evidence supporting several of the suggestions found in the literature for important ideas in proofs. In this way, it appears these suggestions are not in direct competition with one another, but can coalesce to offer a more complete understanding of how mathematicians think about proofs.

We also point out that for these mathematicians, the important ideas were subjective in that they depended on mathematicians' background knowledge. This was most apparent in the case of novel ideas—a strategy that is novel for one mathematician might be familiar to another, and vice versa. Moreover, a mathematician's ability to encapsulate inferences would depend on his or her ability to identify the general method used in that part of the proof. Finally, several mathematicians agreed that their summary would change depending on who it was to be read by, and for what purpose it was intended. For example, some reported that a summary that was intended to be for their own use as class notes would differ from one based on a research paper which contained ideas relevant to their own research. These findings suggest that there is no single "right" summary of a proof, nor is there a single set of important ideas intrinsic to a given proof. Rather, the important ideas seem to depend greatly on who is summarizing the text, for whom, and for what purpose. This point further highlights the difficulty in suggesting that a particular idea represents "the crux" of a proof.

Significant support in our data was found for several ideas suggested by math educators. Raman's (2003) "key idea" could be related to mathematicians' reporting that a diagram could be an important idea in a proof. For Raman, it was the mapping between the diagram and the formal proof that constituted the "key idea". M2 alludes to following the reasoning of a proof on a diagram, which may be a similar notion. Leron's (1983) "high-level ideas" are equivalent to what we have called "goals"—statements in the proof that broadly represent the steps taken in the proof, and which provide a skeleton or structure of the proof. Finally, Rav's (1999) "method" corresponds to our "encapsulation of inferences".

We also note that the students in this study did not mention the important theorems or facts, encapsulation of inferences, or goals of the proof to be important. Moreover, they mentioned the proof technique to be important whereas the mathematicians did not. This suggests that students focus on different aspects of a proof than mathematicians do.

5.4. Limitations of this study

A single exploratory study is insufficient to make generalizable claims about the behavior of students and mathematicians at large. We are only able to make hypotheses regarding the types of important ideas that these populations might consider important. We also note that we measured what participants found to be important by noting what they mentioned in the interviews. It might be the case that someone may have found something to be important but did not mention it during the interview. However, since participants engaged with multiple proofs and were also directly asked what kinds of ideas they found important at the end of the interview, this possibility seems somewhat remote. Finally, our model is grounded in the data presented above. More empirical studies should be conducted to test this model.

5.5. Implications for teaching and research

We also hypothesize that students do not generally summarize proofs that they have written. If this last hypothesis is correct, instructors can remedy this by encouraging students to reflect on the proofs that they have written, and giving them opportunities to revise and summarize proofs that they have turned in.

Further research should be conducted on students' summaries of proofs. As discussed in section 2, reading comprehension research lends promise to the idea that summarization improves comprehension, and is a skill that can be taught. If a similar result holds in mathematics, this could be an easy way to increase students' proof comprehension.

References

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. Journal of Mathematical Behavior, 24, 125-134

Arzarello, F. (2007). The proof in the 20th century: From Hilbert to automatic theorem proving. In P. Boero (ed.), Theorems in schools: From history, epistemology, and cognition to classroom practice. Rotterdam, The Netherlands: Sense Publishers.

Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. Educational Studies in Mathematics, 42, 225-235.

Corbin, J., & Strauss, A. (2008). Basics of qualitative research (3rd Ed.). Los Angeles, CA: Sage. Cunningham, J. (1982). Generating interactions between schemata and text. In New inquiries in reading: Research and instruction, thirty-first yearbook of the National Reading Conference. International Reading Association.

Dole, J., Duffy, G., Roehler, L., & Pearson, P.D. (1991). Moving From the Old to the New: Research on Reading Comprehension Instruction. *Review of Educational Research*, Summer 61: 239-264

Duke, N.K. & Pearson, D. (2002). "Effective Practices for Developing Reading Comprehension." In A.E. Farstrup & S.J. Samuels (Eds.). What Research Has To Say About Reading Instruction(3rd ed., pp. 205–242). Newark, DE: International Reading Association.

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

Hidi, S. & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research, 86,* 473-493.

Leron, U. (1983). Structuring mathematical proofs. American Mathematical Monthly, 90(3), 174-184 McNeil & Donant, 1982

Mejia-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

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.

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.

Raman, M. & Weber, K. (2006). Key ideas and insights in the context of three high school geometry proofs. Mathematics Teacher, 99 (9), 644.

Rav, Y. (1999). Why do we prove theorems? Philosophia Mathematica, 7(3), 5-41.

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.

Thiede, K. W. & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. Contemporary Educational Psychology, 28, 129-160.

Weber, K. (2001). Student difficulties in constructing proofs: The need for strategic knowledge. Educational Studies in Mathematics, 48(1), 101-119.

Weber, K. (2006). Investigating and teaching the thought processes used to construct proofs. Research in Collegiate Mathematics Education, 6. 197-232.

Weber, K. and Alcock, L. (2005). Using warranted implications to read and understand proofs. For the Learning of Mathematics, 25(1), 34-38

Weber, K. and Mejia-Ramos, J.P. (in press). Why and how mathematicians read proofs. To appear in Educational Studies in Mathematics.

Yang, K.-L., & Lin, F.-L. (2008). A model of reading comprehension of geometry proof. Educational Studies in Mathematics, 67, 59-76.

Appendix

Proof 1

Given: D is the midpoint of AB; F is the midpoint of AC; E is the midpoint of BC. (See Figure 6)

Claim: Area(ΔEFD) = ¼Area(ΔABC)

Proof:

Statement	**Reason**
1. AD = DB; AF = FC; BE = EC	Definition of midpoint
2. AB = 2·AD; AC = 2·AF	Line segment addition
3. AB/AC = (2·AD)/(2·AF)	Division property of equality
4. (2·AD)/(2·AF) = AD/AF	Cancellation
5. AB/AC = AD/AF	Transitive property of equality
6. □∠□=□ ∠□	Reflexive property of equality
7. ΔABC is similar to ΔADF	Side-angle-side similarity
8. AB = 2·DB; BC = 2·BE	Line segment addition
9. AB/BC = (2·DB)/(2·BE)	Division property of equality
10. (2·DB)/(2·BE) = DB/BE	Cancellation
11. AB/BC = DB/BE	Transitive property of equality
12. □∠□=□ ∠□	Reflexive property of equality
13. ΔABC is similar to ΔDBE	Side-angle-side similarity
14. ΔABC is similar to ΔFEC	Similar reasoning to the above
15. □∠□=□ ∠□□□;□ ∠□□□=□ ∠□	Corresponding angles of similar triangles are congruent

16. AD = DB	Line 1
17. ΔADF≅ΔDBE	Angle-side-angle congruence
18. □∠□=□ ∠□□□;□ ∠□□□=□ ∠□	Corresponding angles of similar triangles are congruent
19. AF = FC	Line 1
20. ΔADF≅ΔFEC	Angle-side-angle congruence
21. AD = FE	Corresponding parts of congruent triangles are congruent
22. AF = DE	Corresponding parts of congruent triangles are congruent
23. DF = DF	Reflexive property of equality
24. ΔADF ≅ΔEFD	Side-side-side congruence
25. ΔEFD≅ΔADF≅ΔDBE≅ΔFEC	Lines 22, 25, and 29
26. Area(ΔABC) = Area(ΔADF) + Area(ΔDBE) + Area(ΔFEC) + Area(ΔEFD)	Additive property of area
27. Area(ΔABC) = Area(ΔEFD) + Area(ΔEFD) + Area(ΔEFD) + Area(ΔEFD)	The areas of congruent triangles are equal
28. Area(ΔABC) = 4·Area(ΔEFD)	Simplification
29. Area(ΔEFD) = ¼Area(ΔABC)	Division property of equality

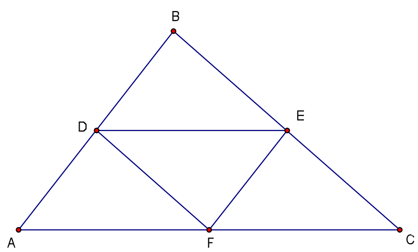


Figure 6: Included with Proof 1

Proof 2

Claim.

The only solution to the equation □*3+5*□*=3* □*2*+sin□ is □*=0*.

Proof.

Clearly, *03−50=302*+sin*0*, so □*=0* is a solution to the equation. We need to show there are no other solutions.

Let □□=□*3−3*□*2+5*□− sin□.

Roots of □□*=0* precisely correspond to solutions of □*3+5*□*=3* □*2*+sin□.

Suppose □□*=0* has a nonzero root; that is □*≠0* and □□*=0*.

□′□*=3*□*2−6*□*+5*− cos□*=3*□*2−2*□*+1* *+2*−cos□*=3*□*−12+2*−cos□.

Since *3*□*−12≥0* and *2*−cos□*>0* for all real numbers *x*, □′□*>0* for all real numbers *x*.

Since □*0*=□ □*=0* and, □*≠0* by Rolle's theorem, there exists *c* between 0 and *s* such that □′□*=0*.

However, this is a contradiction because □′□*>0* for all *x*.

Note: Rolle's theorem states that if *f* is a differentiable function, $a < b$, and $f(a) = f(b)$, then there is a *c* such that □<□<□ and □′□*=0* .

Proof 3

Definition: Two or more natural numbers are *coprime* if the largest natural number which divides all of them is 1.

Definition: A triple (a,b,c) of natural numbers is called a *primitive Pythagorean triple* if $a^2+b^2=c^2$ and a, b, and c are coprime.

Background fact 1: Let $s=p_1^{e_1}\cdot p_2^{e_2}\cdot p_3^{e_3}\cdots p_n^{e_n}$ be the prime decomposition of s. Then s is a perfect square if and only if e_i is even for all i, $1\le i\le n$.

Background fact 2: If k^2 divides m^2 then k divides m.

Claim: If (a,b,c) is a primitive Pythagorean triple, and *a* and *c* are odd, then $\frac{c-a}{2}$ and $\frac{c+a}{2}$ are both perfect squares.

Example: (3,4,5) and (20,21,29) are primitive Pythagorean triples and we have:

$\frac{5-3}{2}=1=1^2$, $\frac{5+3}{2}=4=2^2$, and $\frac{29-21}{2}=4=2^2$, $\frac{29+21}{2}=25=5^2$

Lemma: Suppose s and t are coprime integers and st is a perfect square. Then s and t are perfect squares.

Proof of Lemma: Let $s=p_1^{e_1}\cdot p_2^{e_2}\cdot p_3^{e_3}\cdots p_n^{e_n}$ and $t=q_1^{f_1}\cdot q_2^{f_2}\cdot q_3^{f_3}\cdots q_m^{f_m}$ be the prime decompositions of s and t, where all the exponents are nonzero. Since s and t are coprime, $p_i\neq q_j$ for all i, j, $1\le i\le n$, $1\le j\le m$. Hence, $st=p_1^{e_1}\cdots p_n^{e_n}\cdot q_1^{f_1}\cdots q_m^{f_m}$ is the prime decomposition for st. Since st is a perfect square, by Background Fact 1, above, $e_1,\ldots,e_n$ and $f_1,\ldots f_m$ are even. Hence, s and t are perfect squares.

Proof of Claim:

First, note that since c and a are both odd, $c-a$ and $c+a$ are both even, so $\frac{c-a}{2}$ and $\frac{c+a}{2}$ are integers.

Next, we show that $\frac{c-a}{2}$ and $\frac{c+a}{2}$ are coprime. Let k be an integer dividing $\frac{c-a}{2}$ and $\frac{c+a}{2}$. Then k must divide $\frac{c+a}{2}+\frac{c-a}{2}=c$. Similarly, k must divide $\frac{c+a}{2}-\frac{c-a}{2}=a$. But then k^2 divides a^2 and k^2 divides c^2, so then k^2 divides $b^2=c^2-a^2$. Hence k divides b, by Background Fact 2. Therefore, k divides a, b, and c, so $k=1$ since a, b and c are coprime.

Since $b^2=c^2-a^2=(c-a)(c+a)=4\cdot\frac{(c-a)}{2}\cdot\frac{(c+a)}{2}$, 4 divides b^2. Hence 2 divides b (by Background Fact 2) and b/2 is an integer. Now, $\left(\frac{b}{2}\right)^2=\frac{(c-a)}{2}\cdot\frac{(c+a)}{2}$ and $\frac{(c-a)}{2},\frac{(c+a)}{2}$ are coprime, so by the lemma, $\frac{(c-a)}{2}$ and $\frac{(c+a)}{2}$ are perfect squares.

WHAT DO MATHEMATICIANS DO WHEN THEY HAVE A PROVING IMPASSE?

Milos Savic
New Mexico State University

This paper reports what six mathematicians did when they came to impasses while constructing proofs on an unfamiliar topic, from a set of notes, alone, and with unlimited time. Detailed information is given on two of the mathematicians. By an impasse, I mean a period of time during the proving process when a prover feels or recognizes that his or her argument has not been progressing and that he or she has no new ideas. What matters is not the length of time but its significance to the prover and his or her awareness thereof. I point out two kinds of actions these mathematicians took to recover from their impasses: one kind relates directly to the ongoing argument, while the other kind consists of doing something unrelated, either mathematical or non-mathematical. Data were collected using technology and a new technique being developed to capture individuals' autonomous proof constructions in real-time.

Key words: university level, proof, mathematicians, impasse, data collection technique

This preliminary report presents findings from part of an ongoing larger study of mathematicians, graduate students, and undergraduates constructing proofs on an unfamiliar topic, from a set of notes, alone, and with unlimited time. During separate data collection sessions with nine mathematicians (eight males and one female), six of the nine experienced a considerable period during which they made little progress and developed no new ideas in proving certain theorems. The study investigated what actions these mathematicians took to try to continue, as well as what they later indicated they normally do in such situations. Data were collected employing a new data collection technique being developed to capture individuals' autonomous proof constructions in real-time using a tablet computer or using a LiveScribe pen with special paper. Semi-structured exit interviews were conducted after each mathematician's proving session, followed by focus group reflective interviews conducted with the four professors who used the tablet PC and separately with four of the professors who used a LiveScribe pen (one professor using the LiveScribe pen could not attend the focus group interview). Results, that is, an understanding of how mathematicians recover from periods of no progress and no new ideas, is likely to play a role in facilitating students' learning of proof construction.

Background Literature

Research on mathematicians' practices has been conducted largely in order to guide teaching. Weber (2008) stated that, "investigations into the practices of professional mathematicians should have a strong influence on what is taught in mathematics classrooms" (p. 451). Burton (1999) studied mathematicians specifically for their "attitudes, beliefs and practices…when engaged upon a research problem" (p. 121). Misfeldt (2003) examined how mathematicians make use of writing and re-writing in their own research. Carlson and Bloom (2005) investigated mathematicians' practices during problem solving, an important component of constructing proofs. More recently, Wilkerson-Jerde and Wilensky (2011) explored mathematicians' learning of new mathematics, noting that, "learning new mathematics is an important part of expert practice for professional mathematicians" (p. 23). Samkoff, Lai, and Weber (2011) examined how mathematicians use diagrams to construct proofs.

This paper continues the above research's common thread of examining mathematicians' practices in doing and learning mathematics, in particular, in constructing proofs. Its findings should be useful in teaching proof construction in a way that complements prior research on university students' proving including: difficulties they encounter during the proving process (Moore, 1994; Weber & Alcock, 2004), difficulties with validations of proofs (Selden & Selden, 2003), and difficulties with comprehension of proofs (Conradie & Frith, 2000; Mejia-Ramos, et al., 2010) as well as Harel and Sowder's (1998) categorization of students' proof schemes, that is, the ways they decide what is true.

In analyzing mathematicians' proof construction practices, this paper focuses on impasses, as well as incubation and insight. These ideas have been used in the computer science, psychology, creativity, and mathematics education literatures, mainly in analyzing problem solving. A brief discussion of this literature will provide a background for this paper's usage of the terms in analyzing proof construction. Duncker (1945) defined an impasse as a "mental block against using an object in a new way that is required to solve the problem." He stated that, "[real] problem solving starts when a solver comes to an impasse." In contrast, Van Lehn (1990) appears to mean something different by an impasse. In his work on multi-digit subtraction, he described four categories of impasses in the execution of procedural knowledge. Those impasses were categorized by differences in the actions leading to the impasse, and were treated like "computer bugs."

Some computer scientists concerned with automatic theorem provers have a different meaning for (machine) impasses. Meier and Melis (2006) pointed out that an automatic theorem prover "gets stuck" when the computer has no further techniques with which to solve the current problem and ceases its pursuit of a proof. The actions programmed to attempt to overcome such impasses include building "proof plans," but even such plans have their limitations because "some proofs contain parts that are unique to that proof" (Lowe, Bundy, & McLean, 1998). Meier and Melis (2006) mentioned an advantage that humans have over automatic theorem provers: "When an expected progress does not occur or when the proof process gets stuck, then an intelligent problem solver [i.e., a person] analyzes the failure and attempts a new strategy."

One way human problem solvers sometimes recover from an impasse is through incubation. Incubation, according to Wallas (1926), is the process by which the mind goes about solving a problem, subconsciously and automatically. It is the second of Wallas' four stages of creativity:

- preparation (thoroughly figuring out what the problem is),
- incubation (when the mind goes about solving a problem subconsciously and automatically),
- illumination (receiving an idea after the incubation process), and
- verification (figuring out if the idea is correct).

In later work, Smith and Blankenship (1991) stated that "the time in which the unsolved problem has been put aside refers to the *incubation time*; if [illumination] occurs during this time, the result is referred to as an *incubation effect*" (p. 61). It has been conjectured that this effect happens best when one takes a break from creative work (Krashen, 2001). Illumination is also referred to by some authors as *insight*, or as a "Eureka" or "Aha!" moment (Bowden, Jung-Beemen, Fleck, & Kounios, 2005). Of such moments, Beeftink, van Eerde, and Rutte (2008) observed: "Individuals suddenly and unexpectedly get a good idea that brings them a great step further in solving a problem." In the neuroscience literature, Christoff, Ream, and Gabrieli (2004) have noted that insight, which might appear to be a spontaneously occurring thought process, "share[s] executive and cognitive mechanisms with goal-directed thought".

Both incubation and insight have been studied in psychology with mixed success. According to Smith and Blankenship (1991), "Several empirical studies have tested incubation effects in problem solving. A few of these experiments found incubation effects…In sum, these studies provide neither a strong base of empirical support for the putative phenomenon of incubation nor a reliable means of observing the phenomenon in the laboratory" (p. 62). More recently, Sio and Ormerod (2009), in their meta-analysis of 29 articles covering 117 separate experiments dealing with incubation, concurred: "Although some researchers have reported increased solution rates after an incubation period, others have failed to find effects" (p. 94). Psychologists have tried to provide better ways of capturing the creative process, including reinterpreting several theories of incubation (Helie & Sun, 2010), but have yet to provide consistent concrete evidence of success or failure.

To date the research on incubation in the mathematics education research literature has been sparse and primarily anecdotal. Byers (2007), in his view of creativity in mathematics, described stages similar to those of Wallas. In his investigation of mathematicians' practices, Hadamard (1945) mailed surveys to mathematicians around the world to develop his own ideas of what mathematicians do. More recently, in his dissertation research on AHA! experiences, Liljedahl (2004) used interviews with mathematicians to obtain data on insight. Liljedahl had tried creating an environment for mathematicians to exhibit insight, but conceded: "A further flaw in my experimental design was the role that the environment and setting play in the facilitation of AHA! experiences. Upon reflection, I now see that the clinical interview is not at all conducive to the fostering of such phenomena…" (p. 49). Still, both Hadamard and Liljedahl uncovered some evidence that mathematicians use incubation and then experience insights when solving problems. I hope to add to this literature, partly through narrowing the focus to theorem proving, making observations in a realistic setting, and supplying notes on an unfamiliar topic for the mathematicians to work on.

Theoretical Framework

By an *impasse*, I mean a period of time during the proving process when a prover feels or recognizes that his or her argument has not been progressing fruitfully and that he or she has no new ideas. What matters is not the exact length of time, or the discovery of an error, but the prover's awareness that the argument has not been progressing and requires a new direction or new ideas. Mathematicians themselves often colloquially refer to impasses as "being stuck" or "spinning one's wheels." This is different from simply "changing directions," when a prover decides, without much hesitation, to use a different method, strategy, or key idea, and the argument continues.

There appear to be two main kinds of mental or physical actions that provers take to recover from an impasse. One kind of action relates directly to the ongoing argument. The other kind of action consists of doing something unrelated which can be either mathematical or non-mathematical. Examples of both kinds will be provided below in the "Results" section.

While the treatments of impasses, incubation, and insight mentioned in the section on "Background Literature" may be useful in investigating a wide view of creativity and problem solving, constructing proofs in mathematics seems to be a topic that calls for some modification of the ideas. For example, all of the 117 experiments considered by Sio and Ormerod (2009) in their meta-analysis of incubation studies used an incubation period of just 1-60 minutes, but mathematicians routinely take more time to overcome impasses in their research, and their proofs tend to be rather long and complex. With this in mind, I define *incubation* as a period of time, following an attempt to construct at least part of a proof, during which similar activity does not occur, and after which, an insight (i.e., the generation of a new idea moving the argument forward) occurs. There might be ultimate success or failure with an insight arising from incubation, but that can only be determined by subsequent verification of the new idea's usefulness. A long proving process might entail several impasses and a number of incubation periods (and subsequent insights), only some of which ultimately contribute to the final proof.

New Data Collection Technique

Nine mathematicians (three algebraists, three topologists, two analysts, and one logician) agreed to participate in this study on proving. They were provided with notes on semigroups (Appendix A) containing 10 definitions, 7 requests for examples, 4 questions to answer, and 13 theorems to prove. The notes were a modified version of the semigroups portion of the notes for a Modified Moore Method course for beginning graduate students. This topic was selected because the mathematicians would hopefully find the material easily accessible, and because there are two theorems towards the end of the notes (Theorems 20 and 21 of Appendix A) that have caused substantial difficulties for beginning graduate students. During their exit interviews, two mathematicians offered that the choice of semigroups had been judicious, because they had been able to grasp the definitions and concepts quickly, and because at least one of the theorems had been somewhat challenging to prove. The data collection was split into two groups: four mathematicians writing proofs on tablet PCs, and five mathematicians writing proofs with a LiveScribe pen and special paper.

Tablet PC

With the tablet PC group, I approached each mathematician separately to explain how to use the hardware and the software. I explained how to use the stylus that came with the tablet PC and how to turn the tablet PC around in order to be able to write on it. There were two software programs on the tablet PC that the mathematicians were to work with: CamStudio screen-capturing software and Microsoft OneNote, which was the space on which the mathematicians wrote their proof attempts. The mathematicians each kept the tablet PC for a period of 2-7 days. After the tablet PC was returned, I analyzed the screen captures (resembling small movies in real time) and the mathematicians' proof writing attempts. All proof writing attempts on OneNote were exported as PDFs for analysis. One or two days after this initial analysis of a mathematician's work, I conducted an exit interview, during which I asked about their proofs and proof-writing (Appendix B).

LiveScribe Pen

The LiveScribe pen group consisted of five mathematicians. I approached each of these mathematicians separately to explain how to use the LiveScribe pen and special paper. The LiveScribe pen captures both audio and real-time writing using a camera near end of the ballpoint pen. When one presses on the "record" square at the bottom of the special paper with the pen, the pen goes into audio record mode, which then allows for the real-time capturing of the writing and speaking. The pen can be stopped by a "stop" button, and all proving periods are time-and-date stamped. Uploading the pen data to a computer goes through the LiveScribe software, and I exported each mathematician's collected proving periods together in one PDF file called a "pencast." The mathematicians each kept the LiveScribe pen and paper for a period of 1-10 days. I collected the work of each mathematician, analyzed the data for a period of 1-2 days, and then conducted an exit interview with each of them. The questions for this group of mathematicians were the same as those for the first group. Questions can be viewed in Appendix B.

Transition from Tablet PC to LiveScribe Pen

The switch from tablet PC to LiveScribe pen was done for several reasons. First, the tablet PC cost $900 and up, whereas the LiveScribe pens are just $99 and up. Second, the size of a movie file for a tablet PC screen capture of 16 minutes is one gigabyte, whereas an almost five hour proving session on the LiveScribe pen is just 60 megabytes. Third, the mathematicians were much more comfortable with pen and paper than with the tablet PC and a stylus, because they had to learn how to handle the tablet. Fourth, there were no visual or auditory quality differences between the data collected using the two techniques. This allowed for a smooth transition of data collection techniques to one that I felt was the most comfortable for the participants, and provided all the real-time data collection that I needed.

Summary Data

Four of the nine mathematicians that participated in the study had problems with the technology and thus did not produce "live" data. However, all four provided fixed written data, whether it was with the tablet on OneNote or writing on the LiveScribe paper without audio/video recording. From this data I could still conclude that some mathematicians had impasses because they were candid in writing all of their work, including crossing out failed attempts. The average total work time on the technology was two hours and five minutes. This time was calculated by adding the durations of their actual work, obtained from the date and time stamps. The average time from the first "clocked in" time-and-date stamp until the last "clocked out" time-and-date stamp was 19 hours, 56 minutes. The average number of pages written was slightly under 13. These three statistics allow one to conclude that the mathematicians expended considerable effort on the problems. Six of the nine mathematicians had impasses with one of the last two theorems. Most mathematicians worked through most of the theorems very quickly until they got to those final two theorems. Two of these mathematicians will be discussed in detail in the next section.

Results

Here is a description of an impasse, an incubation, and an insight leading to a proof for two of the mathematicians: Dr. A, an applied analyst and Dr. B, an algebraist. The technology worked for Dr. A and part of his work is described using the time-and-date stamps. For Dr. B, the technology did not work well, but good quality fixed written work and the exit interview data allow some of his work to be presented below in paragraph form.

Dr. A

In proving Theorem 21, "If S is a commutative semigroup with minimal ideal K, then K is a group," Dr. A experienced an impasse, an incubation, and a resulting insight. The following abbreviated, interpreted timeline illustrates this.

3:48 PM	7/13/11	9 min.	At this time Dr. A first attempted a proof of Theorem 21. He stopped and moved on to Question 22.

4:01 PM		16 min.	Continuing later, when he had finished Question 22, Dr. A scrolled up to his first proof attempt. He looked at his answer to Question 22, and at the ten minute mark, erased his first proof attempt. He then scrolled back to his proof of Theorem 20, viewed it for one minute, and wrote “the argument above proves that □ has a multiplicative identity in □.” There was a brief pause, after which he scrolled up to the proof of Theorem 20 again for the final 30 seconds. Proving ended for the day at 4:17.
11:07 AM	7/14/11	11 min.	The next day Dr. A again started attempting to prove Theorem 21. But this time he used a mapping □ that multiplied each element by a fixed □*0* (an idea from his own research). He struggled with some computations until the end of this “clocked in” period.
11:32 AM		5 min.	When he “clocked in” again, Dr. A again worked with the mapping idea and then wrote, “I don’t know how to prove that □ itself is a group. For example, I don’t know how to show that there is an element of □ that fixes □*0*,” acknowledging that he was at an **impasse**.
11:38 AM		23 min.	However, Dr. A continued trying unsuccessfully to use his mapping idea.
12:22 PM		6 min.	When Dr. A “clocked in” again, he continued trying unsuccessfully to use his mapping idea. For example, he wrote, “To prove □ is well-defined, let □□*0*=□, □□*1*=□*2*. Let □ be any other element of □ such that □□*0*=□. Choose any □∈□ s.t. □□=□*2*. Then □□*1*=□□□=□□□□*0*=□□□□*0*=□□□=□□*1*. So □*(*□*)* is determined once □□*0* is determined.”
12:55 PM	7/14/11	5 min.	Later on, when he “clocked in” again, after a 33-minute gap (which might be considered an **incubation** period), Dr. A proved Theorem 21 writing “Proof of theorem: We just need to show that □ itself has no proper subideals. But □ is principally generated, i.e., fix any □*0*∈□ and □*={*□□*0:*□∈□*}* since □ is [a] minimal [ideal]. If □ were a proper ideal of □…” Notice that this idea (an **insight**) for proving Theorem 21 differs from the idea he had tried 33 minutes earlier.

Dr. A indicated in his exit interview where he had had an impasse, noting "One has to show there aren't any sub-ideals of the minimal ideal itself, considered as a semigroup, and that's where I got a little bit stuck." This is because the concept of ideal really depends on the containing semigroup, here □ or □. Dr. A also indicated how he consciously generally recovers from impasses: he prefers to get "un-stuck" by walking around, but distractions caused by his departmental duties also help. That is, he often takes a break from his creative work by purposely doing something unrelated. In this case, Dr. A took several such breaks, but only the last one yielded a new idea.

Dr. B

Dr. B experienced an impasse on the penultimate theorem (Theorem 20), "If *S* is a commutative semigroup with no proper ideals, then *S* is a group." Unfortunately with Dr. B, there were no screen captures, but his written proof attempts were very detailed and the exit interview was very informative. He wrote, "Stuck on [Theorem] 20. It seems you need [to hypothesize] $1 \in$ □, but I can't find a counterexample to show this [that the theorem is false]." Dr. B next moved on to the final theorem (Theorem 21), the one on which Dr. A had had an impasse, proved it correctly, and then crossed out his proof, probably because he had used his as yet unproved Theorem 20. After that, he moved on to the final request for examples (Question 22), explaining in his exit interview, "I moved on because I was stuck [on Theorem 20]...maybe I was going to use one of those examples...I might get more information by going ahead." Dr. B's next approach was to attempt to create counterexamples for Theorem 20. After considering his candidates for counterexamples for some time and being interrupted by taking his family to lunch, Dr. B proved both theorems correctly.

In his exit interview, Dr. B stated that he had developed a belief that had confused him, and thought that he needed to assume that there was an identity element. He also said, "I probably spent 30 minutes to an hour trying to come up with a crazy example. I went to lunch and while I was at lunch, then it occurred to me that I was thinking about it the wrong way. So I went back then and it was quick [using that insight]."

Impasse Recovery

Below are descriptions of the various actions the mathematicians in this study used to recover from impasses, listed according to whether they were directly related to the ongoing argument, or not directly related to it.

Some actions were observed in the proving processes of the mathematicians in the data collected while they worked alone, whereas other actions were first mentioned during the exit interviews and focus group discussions. All of the described actions have exit interview or focus group quotes from the mathematicians explaining them.

Impasse recovery actions that are directly related to the argument

(a) Using methods that occurred earlier in the session: Some of the mathematicians in this study tried to use a proving technique that they had used earlier in the proving session to overcome an impasse.

> "It would be fairly easy to prove…it's likely an argument, kind of like the one I already used…" (Dr. H)

(b) Using prior knowledge from their own research: There were mathematicians in this study who tried to use their own research to overcome an impasse.

> "I'm trying to think if there's anything in the work that I do that...I mean some of the stuff I've done about subspaces of □2(ℝ), umm...there are things called principal shift invariance spaces that the word principal comes into play." (Dr. A)

(c) Using a (mental) database of proving techniques: One of the mathematicians, Dr. F, had a (mental) database of proving techniques in her head.

> "Your brain is randomly running through arguments you've seen in the past… standard techniques that keep running through my head, sort of like downloading a whole bunch at the same time and figuring out which way to go." (Dr. F)

(d) Doing other problems in the problem set and coming back to the impasse: Five of the nine mathematicians in the study approached their proving impasses by moving on to consider the rest of the problems in the notes.

> "I moved on because I was stuck...maybe I was going to use one of those examples ... I might get more information by going ahead." (Dr. B)

(e) Generating examples or counterexamples: Three of the mathematicians in the study attempted to construct counterexamples to some of the theorems when they felt a theorem had not been correctly stated.

> "At first I thought, 'How could I prove this?' And I didn't immediately think of a proof. Then I thought, 'what about a counterexample?' and pretty quickly I came up with a counterexample, of course which turns out not to be right." (Dr. G)

Impasse recovery actions that are unrelated to the argument

(a) Doing other mathematics: Some mathematicians indicated that they might go to another project to help them overcome proving impasses.

> "What I try to do is to keep three projects going…I make them in different areas and different difficulty levels…" (Dr. E)

(b) Walking around: Some mathematicians indicated that sometimes they may choose to walk around to overcome a proving impasse.

> "When I'm stuck, I often feel like taking a break. And indeed, you come back later and certainly for a mathematician you go off on a walk and you think about it." (Dr. G)

(c) Doing tasks unrelated to mathematics: This is the second non-mathematical action unrelated to an impasse. This action was also perhaps the most unusual, and Dr. E seemed slightly embarrassed when he reported the action to me.

> "Yeah I'll do something else, and I'll just do it, and if there's a spot where I get stuck or something, I'll put it down and I'll watch TV, I'll watch the football game, or whatever it is, and then at the commercial I'll think about it and say yeah that'll work…" (Dr. E)

(d) Going to lunch/eating: This action was shown to be effective with Dr. B.

> "So I had spent probably the last 30 minutes to an hour on that time period working on number [Theorem] 20 going in the wrong direction. Ok, so I went to lunch, came back, and while I was at lunch, I wasn't writing or doing things, but I was just standing in line somewhere and it [an insight] occurred to me the…(laughs)…how to solve the problem." (Dr. B)

(e) Sleeping on it: The last action to overcome an impasse seems to be the easiest for a mathematician. Proving can involve mental exhaustion, so resting can help one's exploration for new ideas.

> "It often comes to me in the shower…you know you wake up, and your brain starts working and somehow it [an insight] just comes to me. I've definitely gotten a lot of ideas just waking up and saying "That's how I'm going to do this problem."
> (Dr. F)

The first of the above actions, namely, doing other mathematics, is mathematical, whereas the remaining actions are non-mathematical diversions. Most of the actions that the mathematicians took to overcome their proving impasses were enacted more or less automatically and were not mentioned during their proving sessions. However, the mathematicians did acknowledge those actions during their exit interviews or in the focus group discussions.

Discussion

A majority of the nine mathematicians in this study exhibited impasses and recoveries from those impasses, including some due to incubation. Furthermore, there were a number of instances in which impasses and recoveries, or incubations, might have occurred in a way that could not be observed. For example, all of the mathematicians reported that when they first received the notes they immediately read them to estimate how long the proofs might take, but none started proving right away. In addition, there were periods during the proving sessions when nothing was recorded, and there were also substantial gaps between the "clock in" and "clock out" times during the proving sessions. Furthermore, when the mathematicians next "clocked in" after having left a proof attempt without finishing it, they almost always had a new idea to explore.

In the focus groups, the mathematicians also discussed methods of impasse recovery and what amounts to incubation (that can occur independent of an impasse). They all did this in a relaxed, assured way, not like someone discussing something unfamiliar, but rather like someone discussing beliefs built up over some time. They described a remarkable number of ways of recovering from an impasse. Furthermore, they mentioned benefits that appear to go beyond just restarting an argument.

During one focus group interview, Dr. G stated, "When we are working on something, we are usually scribbling down on paper. When you go take a break,… you are thinking about it in your head without any visual aids….[walking around] forces me to think about it from a different point of view, and try different ways of thinking about it, often global, structural points of view." There is no "scribbling on paper." Doing this, he believed, might assist in understanding the structure of a problem or even of an area of mathematics. In a somewhat similar vein, Dr. F offered the following, "You just come back with a fresh mind. [Before that] you're zoomed in too much and you can't see anything around it anymore." This seems to be a somewhat more local broadening perspective.

From Dr. G, one sees that there might still be conscious thought about the current mathematical problem going on during a break so he is not referring just to incubation. Dr. F added that "freshness" of mind might also help with overcoming proving impasses. Also, simply going away from and coming back to a problem or proof might yield new ideas for recovering from an impasse. Dr. A stated, "I do have a belief that if I walk away from something and come back it's more likely that I'll have an idea than if I just sit there." These remarks indicate that some mathematicians take deliberate actions to overcome impasses and also to improve the breadth or quality of their perspectives.

Conscious, or deliberate, incubation has been shown in the psychology literature to result in a greater incubation effect than merely being interrupted during the problem-solving process. "Individuals who took breaks at their own discretion (a) solved more problems and (b) reached fewer impasses than interrupted individuals" (Beeftink, van Eerde, & Rutte, 2008). Ironically, interruption seems to have been useful in the case of Dr. B, who said that he would have worked non-stop if he had not been interrupted for lunch with his family. This also agrees with the psychology literature: “It was also found that interrupted individuals reached fewer impasses than individuals who worked continuously on problems” (Beeftink, van Eerde, & Rutte, 2008).

Educational Implications

The Results and Discussion sections above suggest that proving impasses, recoveries from them, incubation, insight, and the ability to deal with such topics is a significant part of doing mathematics, and in particular, of constructing proofs. Thus, it is worth examining how they might be taught. The ways of doing this are yet to be examined in detail, however, one small example can be provided. The professors in this study were unaware of the origin of the notes (Appendix A), and one tried to construct counterexamples. In fact, the notes were designed for teaching beginning graduate students about proving. Theorems 20 and 21 can be made much easier by adding a comment about careful reading of the definition of ideal (Definition B of Appendix A) and by adding two easily proved lemmas for Theorem 20. These were omitted from the notes to provide students with experiences similar to those of these professors’. Most students would probably require several attempts and some advice for proving Theorems 20 and 21. However, the experience of trying may still be valuable.

Similar experiences can probably be provided to students who are not yet familiar with constructing proofs by considering problem solving. A problem that is likely to generate impasses is probably close to what Schoenfeld (1982) described as a “rich” problem:

- The problem needs to be accessible. That is, it is easily understood, and does not require specific knowledge to get into.
- The problem can be approached from a number of different ways.
- The problem should serve as an introduction to important mathematical ideas.
- The problem should serve as a starting point for rich mathematical exploration and lead to more good problems (as cited by Liljedahl, 2004, pp. 187-188).

Notice that in the list of actions to overcome impasses, the mathematicians moved on to consider the request for examples (Question 22), having observed that considering them might be useful. This action to overcome an impasse relates well to students’ experiences, because homework assignments usually consist of multiple problems, so they can go ahead to another problem when they are “stuck.” Furthermore, students may need to experience successes in order to acquire confidence in their proving ability, and telling them what mathematicians do when they “get stuck” might help them when they have “no idea what to do next.” Moreover, there is encouragement from the psychology literature about the positive effects of incubation in the classroom. Sio and Ormerod (2009) listed four articles where “educational researchers have tried to introduce incubation periods in classroom activity, and positive incubation effects in fostering students’ creativity have been reported.” (p. 94)

Future Research

Using LiveScribe pens and the corresponding paper provides a naturalistic setting for provers while gathering real-time data from them. If one can see what a mathematician does during the proving process, those same techniques might be used with students in a transition-to-proof or proof-based course. How can we use this data collection technique in the classroom? Will it benefit students to have LiveScribe pens with which to do their homework so that teachers can analyze their proving processes?

How can we gain additional information on when and how incubation is used in mathematics by mathematicians or students? How can we collect more of the actions that mathematicians use to recover from impasses? Also, how can we encourage students to take some of these actions to recover from their proving and problem-solving impasses?

References

Beeftink, F., van Eerde, W., & Rutte, C. (2008). The effect of interruptions and breaks on insight and impasses: Do you need a break right now? *Creativity Research Journal, 20*, 358-364.

Bowden, E., Jung-Beemen, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *TRENDS in Cognitive Science, 9*, 322-328.

Burton, L. (1999). The practices of mathematicians: What do they tell us about coming to know mathematics? *Educational Studies in Mathematics, 37*, 121-143.

Byers, W. (2007). *How mathematicians think: Using ambiguity, contradiction, and paradox to create mathematics.* Princeton, NJ: Princeton University Press.

Carlson, M., & Bloom, I. (2005). The cyclic nature of problem solving. *Educational Studies in Mathematics 58*, 45-75.

Christoff, K., Ream, J., & Gabrieli, J. (2004). Cognitive and neural basis of spontaneous thought processes. *Cortex, 40*, 623-630.

Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics, 42*, 225-235.

Duncker, K. (1945). *On problem solving.* Washington, D.C.: American Psychological Association.

Hadamard, J. (1945). *The Mathematician's Mind.* Princeton: Princeton University Press.

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

Helie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychology Review, 117*, 994-1024.

Krashen, S. (2001). Incubation: A neglected aspect of the composing process? *ESL Journal, 4*, 10-11.

Liljedahl, P. (2004). *The AHA! experience: Mathematical contents, pedagogical implications.* Vancouver: Simon Frasier University.

Lowe, H., Bundy, A., & McLean, D. (1998). The use of proof planning for cooperative theorem proving. *Journal of Symbolic Computation, 25*, 239-261.

Meier, A., & Melis, E. (2006). Impasse-driven reasoning in proof planning. In M. Kohlhase (Ed.), *Mathematical Knowledge Management: 4th International Conference MKM 2005* (pp. 143-158). Berlin: Springer-Verlag.

Mejia-Ramos, J. P., Weber, K., Fuller, E., Samkoff, A., Search, R., & Rhoads, K. (2010). Modeling the comprehension of proof in undergraduate mathematics. *Proceedings of the 13th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 1-22). Raleigh, NC.: Available online.

Misfeldt, M. (2003). Mathematicians' writing. *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education* (pp. 301-308). Cape Town, South Africa: International Group for the Psychology of Mathematics Education.

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

Samkoff, A., Lai, Y., & Weber, K. (2011). How mathematicians use diagrams to construct proofs. *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 430-444). Portland, OR: Available Online.

Schoenfeld, A. H. (1982). Measures of problem-solving performance and of problem-solving instruction. *Journal for Research in Mathematics Education, 13*, 31-49.

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.

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 35*, 94-120.

Smith, S., & Blankenship, S. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology, 104*, 61-87.

VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions.* Cambridge, MA: MIT Press.

Wallas, G. (1926). *The art of thought.* New York: Harcourt Brace.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 30*, 431-459.

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*, 209-234.

Wilkerson-Jerde, M. H., & Wilensky, U. J. (2011). How do mathematicians learn math? Resources and acts for constructing and understanding mathematics. *Educational Studies in Mathematics, 78*, 21-43.

Appendix A

Definition A: A semigroup *(□,·)* is a nonempty set □ together with a binary operation · on □ such that the operation is associative. That is, for all □,□*,* and □∈□, □·□·□=□·□·□.
Note: We often refer to the "semigroup □" instead of the "semigroup *(□,·)*" and symbols such as +,−,∗,⊙,⊕ may be used instead of "·". Also, ″□·□" or "□□" is often read "□ times □".

Example 1: Find several examples of semigroups.

Definition B: A nonempty subset □ of a semigroup □ is called a left ideal [right ideal, ideal] of □ if □□⊆□ □□⊆□*,* □□∪□□⊆□ where □□=*{*□□|□∈□ and □∈□*}*.

Example 2: Find some examples of left ideals, right ideals, and ideals in several semigroups.

Theorem 3: The intersection of a left ideal and a right ideal is nonempty.

Theorem 4: The intersection of two ideals is an ideal.

Definition C: A non-empty subset □ of a semigroup □ is called a subsemigroup of □ if □□⊆□.

Note: In a semigroup, every left ideal, right ideal, and ideal is a subsemigroup.

Definition D: A semigroup □ is called commutative or Abelian if, for each □ and □∈□, □□=□□.

Example 5: Find some ideals in *[0,1]* under multiplication.

Definition E: An element □ of a semigroup □ is called an idempotent if □□=□. ("□□" is often written "□*2*".)

Definition F: An element *1* of a semigroup □ is called an identity element of □ if, for each □∈□*,* *1*□=□*1*=□. (Other symbols, such as "□", may be used instead of "*1*" to represent an identity element.)

Definition G: An element *0* of a semigroup □ is called a zero element of □ if, for each □∈□*,* *0*□=□*0*=*0*. (Other symbols may be used instead of "*0*" to represent a zero element.)

Example 6: Find a semigroup with an idempotent which is neither the identity nor a zero.

Definition H: An ideal [left ideal, right ideal] □ of a semigroup □ which does not properly contain any other ideal [left ideal, right ideal] of □ is called a minimal [left, right] ideal of □.

Example 7: Find some semigroups that contain, and some that do not contain, a minimal ideal.

Question 8: Can a semigroup be its own minimal ideal?

Theorem 9: Every semigroup has at most one minimal ideal.

Example 10: Find examples of semigroups that (1) are not commutative, (2) do not have idempotents, and (3) consist entirely of idempotents.

Theorem 11: A semigroup can have at most one identity element and at most one zero element.

Theorem 12: Distinct minimal left [right] ideals of a semigroup are disjoint.
Note: If a semigroup has a minimal ideal, it is unique (by Theorem 9) and it is called the kernel of the semigroup. The theory of semigroups started (in 1928) when Suschewitsch characterized the kernel.

Definition I: Let □ and □ be semigroups and □*:*□→□ be a function. We call □ a homomorpism if, for each □∈□ and □∈□, □□□=□□□*(*□*)*. If □ is also one-to-one, □ is called an isomorphism. We say □ and □ are isomorphic if □ is an onto isomorphism.

Note: We think of semigroups □ and □ as the "same" if there is an onto isomorphism □*:*□→□.

Example 13: Find some examples of homomorphisms that are, and are not, isomorphisms. Also find some examples that are, and are not, onto.

Theorem 14: Let □ and □ be semigroups and □*:*□→□ be a homomorphism. If □∈□ is an idempotent, then □*(*□*)* is an idempotent.

Theorem 15: Let □ and □ be semigroups and □*:*□→□ be a homomorphism. If □ is a subsemigroup of □, then □*(*□*)* is a subsemigroup of □.

Theorem 16: Let □ and □ be semigroups and □*:*□→□ be an onto homomorphism. If □ is an identity [zero] of □, then □*(*□*)* is an identity [zero] of □.

Theorem 17: Let □ and □ be semigroups and □*:*□→□ be an onto homomorphism. If □ is an ideal of □, then □*(*□*)* is an ideal of □.

Definition J: A semigroup □ is called a group if □ has an identity *1* and if for each □∈□ there is a □*′*∈□ such that □□*′*=□*′*□*=1*.

Theorem 18: Let □ be a group with identity *1*. If □*,*□*′,*□*″*∈□ with □□*′*=□*′*□*=1* and □□*″*=□*″*□*=1*, then □*′*=□*″*. (That is, the element □*′* so that □□*′*=□*′*□*=1* is unique. The element □*′* is called the inverse of □ in □ and written □*−1*.)

Theorem 19: A group has no proper left ideals [right ideals, ideals].

Theorem 20: If □ is a commutative semigroup with no proper ideals, then □ is a group.

Theorem 21: If □ is a commutative semigroup with a minimal ideal □, then □ is a group.

Question 22: For each of parts a, b, and c are the two semigroups isomorphic? Prove you are right.

(a) $(\mathbb{Z},+)$ where $\mathbb{Z}$ is the integers and + is ordinary addition. $2\mathbb{Z},+$ where $2\mathbb{Z}$ is the even integers and + is ordinary addition.

(b) $(\mathbb{R},+)$ where $\mathbb{R}$ is the real numbers and + is ordinary addition. $(0,\infty,\cdot)$ where $(0,\infty)$ is the positive real numbers and $\cdot$ is ordinary multiplication.

(c) (□, ·) where □=$\{0,1,2,3,4\}$ and for □,□∈□, □ · □=□. $(\mathbb{Z}5,\cdot)$ where $\mathbb{Z}5=\{0,1,2,3,4\}$ and □ · □ means "□□ □□□ 5", i.e., ordinary multiplication minus (whole) multiples of 5. For example, $4\cdot4=16-3\times5=1$, $3\cdot4=12-2\times5=2$, and $3\cdot3=9-5=4$, but $2\cdot2=4$.

Appendix B

Interview Questions

1. Was there anything that was particularly difficult or took you long?
2. (When there were delays) What were you thinking of at this point in time?
3. What made you think of _____ (e.g., stabilizer)?
4. What difficulties were there with the technology?
5. (With a very long delay, e.g., of several hours) What did you do in that time period? Did you think about the notes or some theorem in the notes?

Focus Group Questions

1. (Question to get them comfortable) What did you think of these notes?
2. Compare and contrast your experiences with the last 2 theorems.
3. If and when you did get stuck with these notes, how did you handle that?
4. In general, what do you do when you get stuck (in a problem, proof, with your research)?
5. Is there anything else you do or think about when attempting to prove theorems?

EXAMINING STUDENTS' MATHEMATICAL TRANSITION BETWEEN SECONDARY SCHOOL AND UNIVERSITY: WORKING WITH VECTORS AND LINEAR INDEPENDENCE

Natalie E. Selinski
Universität Kassel

To understand the mathematical transition students make between secondary school and the university requires an in-depth look at the mathematical topics students learn at the time of this transition and the contextual, institutional changes that simultaneously occur. This report explores how linear algebra students at both the secondary school and university in Germany understand vectors and linear independence and dependence in the course of video-recorded, think-aloud problem-solving interviews. Analysis of these interviews indicates not only differences in mathematical content and sophistication between secondary school and university students, but also in students' disposition, particularly towards new mathematical experiences. A look at more informal data about the various institutional environments, the Gymnasium and the University, provides a potential reason for these differences. This report concludes with a discussion on how to create a blended analysis of these individual understandings and dispositions and their relationship with the institutional context as a better means of understanding the transition to university-level mathematics.

Key words: [Transition to university mathematics, Linear algebra, Institutional environments, Student disposition]

The gap between secondary school mathematics and university mathematics has proved to be a particularly difficult challenge for students (cf. De Guzmán, Hodgson, Robert, & Villani, 1998; Tall, 1991). To understand the mathematical transition students make between secondary school and the university requires an in-depth look at the mathematical topics students learn at the time of this transition and the contextual, institutional changes that simultaneously occur. In particular, there are certain courses that fall exactly during this transition. In Germany, linear algebra is one such course, with foundational linear algebra topics like vectors and linear independence being introduced in the last years of secondary school then revisited and built upon in the first year at the university.

This study begins by asking how do students think about and work with the ideas of vectors, linear independence, and linear dependence at the secondary school and university and what differences these two distinct groups of students have in viewing and working with these concepts.

The results of this analysis suggest differences not only in how these distinct groups view these concepts, but also in how students approach tasks that require the students to work with these concepts in novel or more unfamiliar settings and their disposition towards these new mathematical situations. This begs the question: how do we account for the differences between the secondary school students and university students? The study conjectures that these differences come from not only the level and sophistication of the mathematical content of their courses, but also from the differences in the institutional settings. The analysis will include a discussion on creating a blended analysis of these individual understandings and dispositions and their relationship with the institutional context as a useful means for understanding the transition to university-level mathematics.

Motivation for the Study

This report comes out of a larger year-long project sponsored by the Fulbright Program studying the transition from the secondary school to the university in Germany. It is critical to understand the differences and similarities between the German education system and the US education system in order to contextualize the motivation for this study, the design of the research and the results and analysis presented in this report.

Figure 1 depicts a summary of the German and US educational systems in the last years of secondary school and the initial years at the university and the courses students encounter at these years in their education.

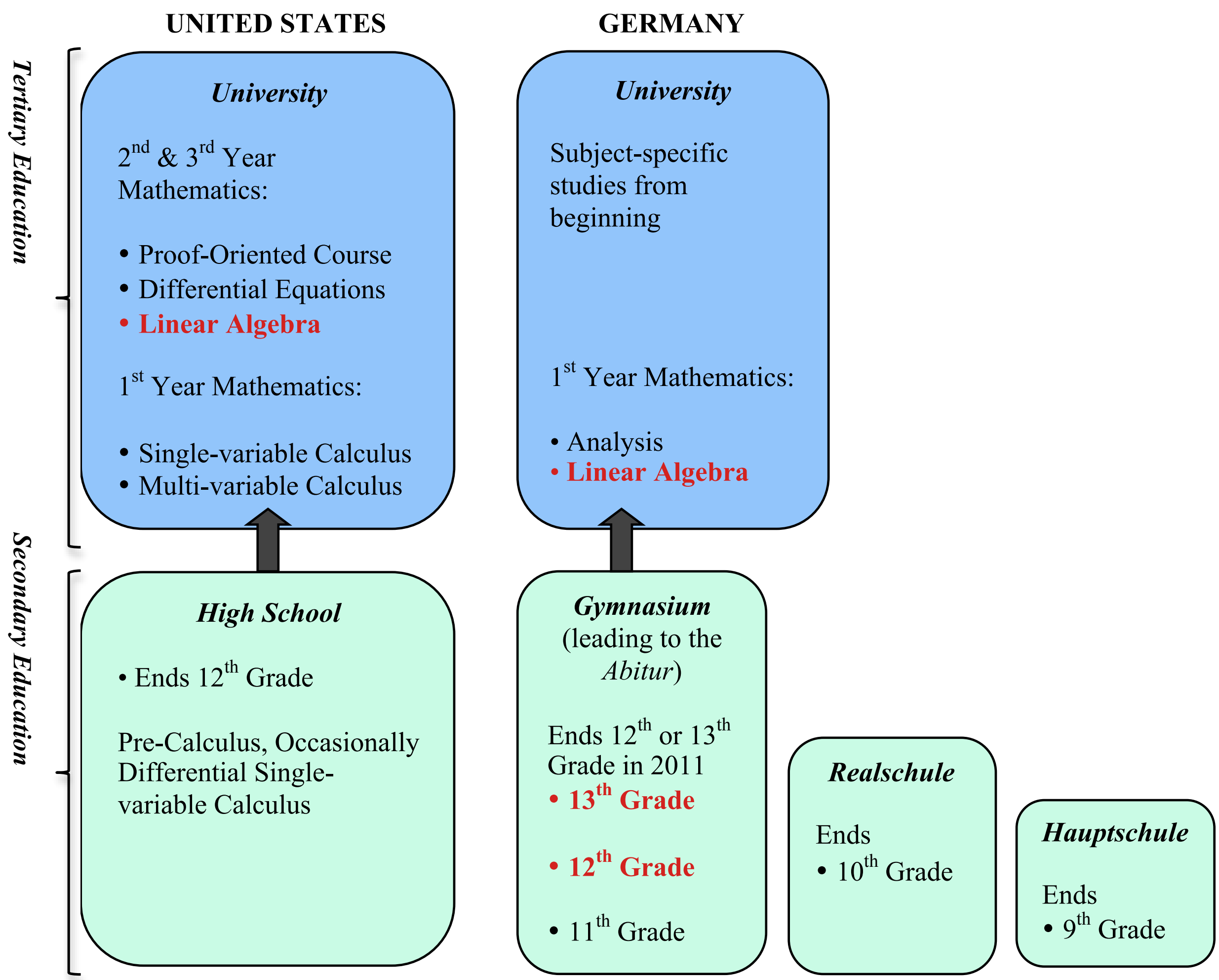


Figure 1. Outline of the US and German education systems, last years of secondary school to the University.

There are multiple differences to observe. The first of which is that the secondary school system in Germany favors a three-tracked system as compared to the US system's favoring of a single-track system. In 2010 about 78% of secondary school students in Germany were in the three-track system (The Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (Standing Conference), 2012). The three-track system consists of three options for secondary schooling: the Hauptschule, the Realschule, and the Gymnasium. As the Standing Conference describes, the

Hauptschule is a lower secondary level school providing basic general education up to 9th grade, and it serves about 16.6% of the secondary school student population in Germany. The Realschule is another option running until 10th grade, providing at the lower secondary level basic general education and at the upper secondary level opportunities for vocational and higher education entrance qualifications, most often at a Hochschule. Of the secondary school student population in Germany, 25% attend a Realschule. The Gymnasium is the typically considered the most advanced secondary school option. The Standing Conference described the Gymnasium as providing in-depth general education aimed at the general higher education entrance qualification through 12th or 13th grade. 36.3% of secondary school students in Germany attend a Gymnasium. Traditionally Gymnasium students complete their secondary studies with a test to obtain the Abitur, a qualification permitting students to study at any institution of higher education in Germany. Furthermore, while students from three tracks may go on to some form of tertiary education, it is traditionally only Gymnasium students who are permitted to then enroll in the University, where students immediately begin their rigorous, more theoretical subject-specific studies.

One result of this three-track system is that the students enrolled in the Gymnasium and, from these Gymnasium students the University students who continue on to study mathematics-related courses at the University, often are exposed to and study mathematics at a faster pace than their American peers. This is particularly prevalent in the case of linear algebra, where students first encounter ideas involving vectors, linear independence and dependence, and some linear algebra proofs both in the 12th and 13th grades mathematics courses, then again in a rigorous first year, year-long linear algebra course. Compare this to linear algebra studies in the US, where most students do not encounter linear algebra until a 2nd or 3rd year at the university.

These drastic differences give rise to a number of questions. What exactly are the differences and similarities in terms of a student's mathematics education between the US and German systems? How are students in Germany able to encounter these ideas so much faster in their mathematical educations? Do German students encounter similar problems to their US counterparts? How do students conceptualize ideas presented first at the Gymnasium, then again at the University? What can we learn about the transition to university-level mathematics by studying students in the German system, particularly when learning mathematical theories and topics that occur as students transition into university-level mathematics? While these questions are too lofty and complex to be addressed in a year-long project, they provide the foundation for my broader research project. Furthermore, these questions provide the background to the more specific study detailed in this report.

Research Questions and Literature Review

The research presented in this report focuses in on much more specific questions within the broader inquiries previously stated. Specifically, the research questions include: How do students think about and work with vectors and linear independence and linear dependence at the Gymnasium level and the University level? How do these ways of thinking about and working with vectors and linear independence and linear dependence exemplify the transition from the Gymnasium to the University? As the study progressed, differences between the students at the Gymnasium and the University emerged, extending beyond differences in mathematical content. This led me to ask what specifically are the differences between these groups of students and how can one account for these differences?

In terms of how students conceptualize and work with concepts in linear algebra, there is a growing body of work detailing various ways in which students learn and think about linear algebra and the struggles students have with linear algebra. One of the most comprehensive collections of research in linear algebra is Dorier's (2000) compilation. Within this book Sierpinska (2000) details certain aspects of students' reasoning in linear algebra, arguing that students favor thinking practically as opposed to theoretically. As a result, students develop "the obstacle of formalism", where they treat the formal symbolic representations of objects in linear algebra as the objects themselves without understanding the structure of these representations. She goes on to describe three modes of reasoning in linear algebra – synthetic-geometric, analytic-arithmetic, and analytic-structural – and the tensions and challenges students experience in using these three historical significant modes.

Within the same compilation Hillel (2000) further examines the difficulties students have when learning linear algebra. His study begins by observing that "linear algebra has traditionally been the first mathematics course that students encounter which is a full-fledged mathematical theory" (p. 191) and this creates many of the difficulties students have. Furthermore, Hillel notes:

> This phenomenon is not as local as one might expect. So, for example, while the teaching of mathematics in France has been a lot more formal than its North American counterpart… French students of linear algebra do not seem to have any easier time with proofs. (p.192)

As the teaching and education system for mathematics in France and in Germany is quite similar, his argument for studying the difficulties in linear algebra as a more global phenomenon is particularly salient for this report.

Hillel goes on to explore three modes of description for basic objects in linear algebra: geometric, algebraic and abstract. The geometric mode consists of using the language and concept of 2- and 3-space, such as describing vector as an arrow with direction and magnitude. The algebraic mode uses the language and concepts of the more specific theory of R^n. A depiction of a vector as an *n*-tuple would belong to the algebraic mode. The abstract mode uses the language and concepts of the general formalized theory, such as depicting a vector as an element of a vector space as defined by a set of axioms. Hillel then explores the difficulties seen in North American students in using each of these modes of description and in transitioning between these modes. While these modes of description and the problems associated with them are naturally linked to Sierpinska's modes of reasoning, these frameworks are distinct. For the purpose of this study, Hillel's modes of description provided a natural tool for unpacking students thinking with vectors and sets of vectors, and thus I will adopt the same modes of description and definitions outlined by Hillel.

Studies have also been conducted regarding student understanding of linear independence and dependence. Bogomolny (2007) examines students' reasoning of linear independence and dependence through example-generation. Specifically, she asks students to generate example 3x3 matrices whose columns are linearly dependent/independent, then to explain their reasoning, the relationship to span, and the way in which these examples were generated. Her analysis highlights both the significance of example-generation as a tool for researchers to understand student thinking and for instructors to contribute to the learning process.

This report also uses example-generation as a means to understanding students' ways of thinking about concepts in linear algebra. Examples are known to play a significant role in the way students learn mathematics. Previous studies suggest that students often rely on worked examples to develop both weak and strong concept-understandings (Mason & Watson, 2008; Weber, Porter, & Housman, 2008). Other research suggests student concept-understanding is richest when students are asked to play an active role in constructing and analyzing examples, thus asking them to understand the underlying structure of the mathematics (Dahlberg & Housman, 1997; Mason & Watson, 2008). This study uses example generation not in a pedagogical way suggested in these studies but as a tool for ascertaining students' understandings of vectors, linearly independent and dependent set of vectors, and the underlying structure of linear algebra.

To further analyze results and the reciprocal nature between the students' individual thinking and understandings and the institutional environments in which they learn, Cobb and Yackel (1996) created an elaboration of the interpretive framework seen in Figure 2.

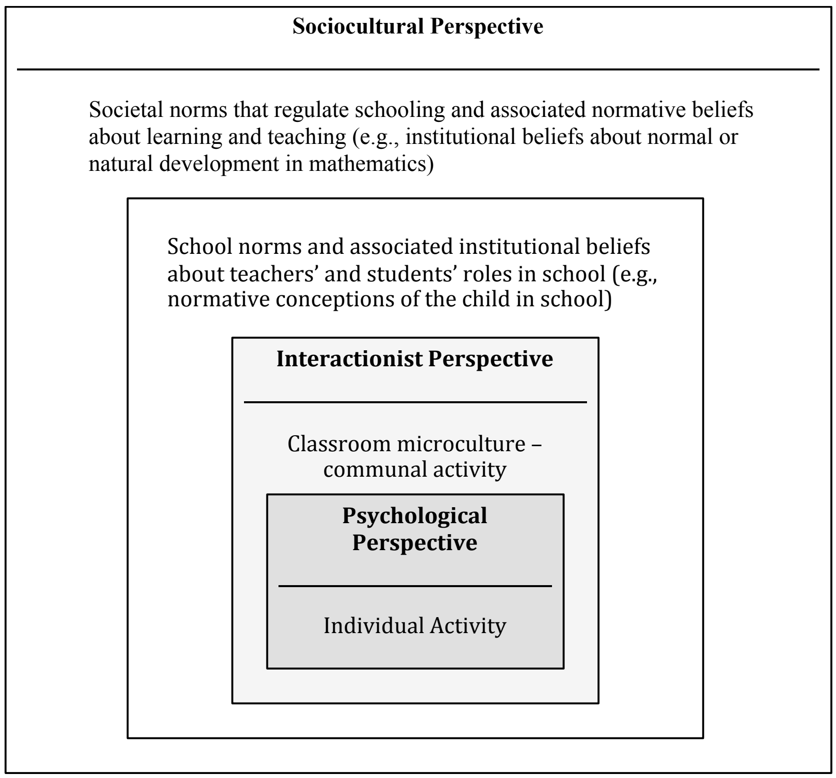


Figure 2. An elaboration of the interpretive framework (Cobb & Yackel, 1996).

This framework was developed to take into account the broader institutional contexts in which psychological constructivist analyses of individual activity and interactionist analyses of classroom interactions and discourse are embedded. While this report does not do an in-depth analysis on the classroom social norms, sociomathematical norms or institutional norms as is essential to use this framework to its fullest potential, it did offer a framework to begin a discussion on the interplay of students' individual thinking and activity and the institutional norms of the broader contexts in which the individual activities are located.

Methodology

Data for this report comes from a year-long project examining how students learn linear algebra at the Gymnasium and University. Data collection came primarily from two sources, course observation and interviews.

The researcher observed complete linear algebra units at the Gymnasium and University. For the Gymnasium, this consisted of attending five 90 minute long lessons regarding vectors and linear independence and dependence, held at most twice per week. Field notes were taken during the lessons, all texts and homework were reviewed by the researcher, and occasional informal, unstructured conversations occurred with the instructor and students to better understand the institutional environment.

At the University, observation consisted of attending a semester-long course on linear algebra, which consisted of two 90 minute long lectures per week, weekly practice sessions with a teaching assistant, and voluntary, informal "Math Room" use. Furthermore, all written notes and homework used for the course were also reviewed by the researcher.

Six Gymnasium students and five University students participated in individual, semi-structured, think-aloud problem-solving interviews (Bernard, 1988) that were approximately 60 to 90 minutes long. It should also be noted that interviews were conducted primarily in German, with the help of a translator who has specific knowledge of mathematics, mathematical vocabulary and teaching mathematics in both Germany and the US. The interviews were video-recorded, and the analysis of the data involves repeatedly reviewing these videos, selective transcriptions of the videos, and copies of students written work created during the course of the interviews. When transcribed portions of the interviews were used, translations were created in collaboration between the researcher and the translator to assure the closest translation possible to the students' utterances. This report will focus on the two questions posed in these interviews, presented in Figure 3.

1. How do you think about what a vector is?
   (Follow-up questions: Do you have a geometric/ algebraic/ abstract understanding? Are these understandings are connected, and if so, how?)

2. For each of the following, please create an example that fits the given criteria:
   a. A set of vectors in R^2 (R^3, R^4, R^n, not in R^n) that is linearly dependent.
   b. A set of vectors in R^2 (R^3, R^4, R^n, not in R^n) that is linearly independent.

   (Follow-up questions: What does it mean for a set of vectors (in R^2, R^3, R^4, R^n, not in R^n) to be linearly dependent? Linearly independent? Do you have a geometric/ algebraic/ abstract understanding? Are these understandings connected, and if so, how?)

Figure 3. Relevant interview questions.

Students were not given all the interview questions at once. Rather, questions were revealed one at a time, and only after completing one part was the next part of the question revealed. Specifically, for Question 2 students were first asked to generate an example of a set of vectors in R^2 that is linearly dependent. Only after the discussion for this question was complete was the next prompt revealed: generate an example of a set of vectors in R^2 that is linearly independent. Students were then asked the same questions to generate examples in R^3, subsequently R^4, R^n, and not in R^n. This permitted the students and researcher to focus on the student's understanding of vectors and linear independence and dependence in each space more clearly.

Follow up questions were frequently asked to encourage students to think aloud, unpack a student's understanding, ask for other modes of description and clarify connections between modes that were utilized. These questions were written as part of the protocol, and thus were asked consistently when students did not readily explain their thinking completely.

These questions were based upon the material both Gymnasium and University students encounter in their linear algebra studies. Furthermore, for the second question, students were asked to generate examples in spaces that were both familiar to them, such as R^2 and R^3, and more novel, such as R^4 for the Gymnasium students and not in R^n for the University students.

Results

Individual Student Problem-Solving Interviews

In the results that follow, it should be noted that two aspects of the data and analysis are emphasized. First, a great deal of discussion will be spent on the modes of description the students from both the Gymnasium and the University favored for each prompt, and secondly, the differences that came out of particularly the later parts of the second interview question, in which the students were asked to generate examples in more novel spaces. While the analysis presented here using modes of description can be extended to students understanding of linear independence and dependence, these do not give rise to results that extend much beyond Hillel's previous work. As such, the aspects of the interviews and analysis that were more surprising are set up and discussed, while a detailed discussion on students' understanding of linear independence and dependence is limited.

All students, both from the Gymnasium and the University, readily depicted vectors using geometric and algebraic modes of description. Johannes, a Gymnasium student, provided a prototypical example in Figure 4.

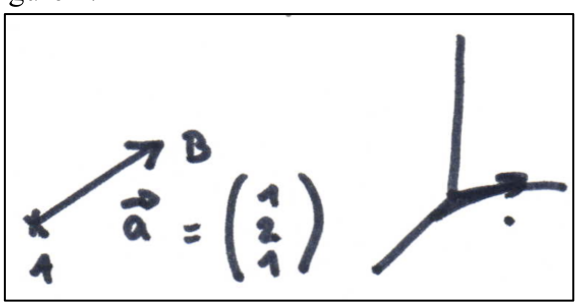


Figure 4. Johannes, a Gymnasium student, provides prototypical geometric and algebraic descriptions of vectors

Initially he described geometrically a free vector, with a set magnitude and direction. When prompted for an example of a vector described algebraically, he created an example of the 3-tuple. Connecting these two representations posed no challenge for Johannes, who explained, "if you have a coordinate system, the three axes, then it goes like 1 in x_1, 2 in x_2, 1 in x_3" (translated) and he drew the vector from the origin to the point (1, 2, 1). This confidence in working with and connect geometric and algebraic modes of description was shown by all students in the study. Given that geometric and algebraic modes were taught and emphasized in a variety of exercises for all students at some point early in their introduction to linear algebra, these favored modes came as no surprise.

What was more surprising was the lack of abstract description given by the students from the University. Despite a semester-long course in linear algebra that favored abstract representations and presented vectors from the very beginning as abstract entities, only one student, Michael, initially showed a more flexible abstract understanding: “If you want to be more abstract, you could say it [a vector] is an element of a vector space.” Michael then proceeded to explain the axioms of a vector space and later used this abstract description of vectors to generate examples for later prompts.

As the interview progressed, most Gymnasium students continued to utilize algebraic and geometric modes, both separate and in combination, to generate examples of linearly dependent and independent sets of vectors in R^2 and R^3. Figure 5 gives Samuel’s examples for linearly dependent and independent sets of vectors in R^3 as well as one definition of what it means for a set of vectors to be linearly independent.

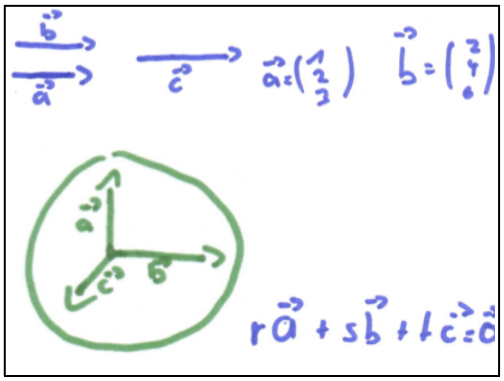


Figure 5. Samuel, a Gymnasium student, uses algebraic and geometric modes of description to generate examples of linearly dependent (top) and independent (bottom) sets of vectors in R^2 and R^3.

However, the Gymnasium students reacted differently when prompted to generate examples in R^4. Three out of the six students interviewed immediately dismissed the possibility of such a space. Rachel most succinctly explained the difficulty cited by all three students: “That [R^4] does not exist…. You already have length, width and height, then that is enough to express everything” (translated). That is, since all physical objects that we see and physically interact with are described with three dimensions, there physically can be no more dimensions in which to generate examples. Though all three students showed the mastery of the algebraic mode of description to generate examples in this novel space, this algebraic mode of description was taught based upon a geometric understanding of vectors. One might conjecture that when this geometric description no longer had meaning, the students did not see meaning in other modes.

For the three Gymnasium students who did not immediately dismiss R^4, there were two approaches. As Elena shows in Figure 6, a purely algebraic mode of description could be used. Elena readily produced vectors as 4-tuples to expand her understanding to R^4. This permitted her to create first a set of linearly dependent vectors in R^4, then a set of linearly independent vectors in R^4.

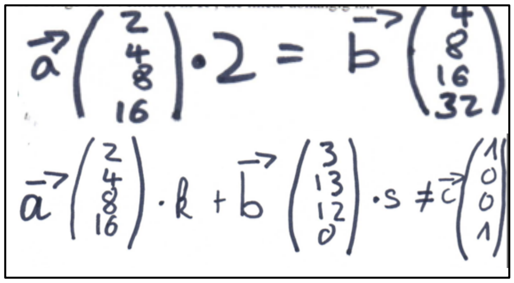


Figure 6. Elena, a Gymnasium student, uses an algebraic mode of description to generate an examples of a linearly dependent (top) and dependent sets of vectors in R^4.

A second approach was given by Johannes. As shown in the quotation that follows, he was able to utilize a more abstract mode to expand his understanding to R^4.

> In 2-dimensional space, there are at most two vectors that can be linearly independent, since the third makes them linearly dependent. In 3-dimensional space, we had at most three vectors that were linearly independent. The fourth made them dependent. In 4-dimensional space, it is exactly like before. At most four vectors can be linearly independent, and the fifth makes them all together dependent. (translated)

Johannes was able to leverage his understanding of the structure of R^2 and R^3, specifically the possible number of linearly independent vectors, to make conjectures about linear independence in R^4. It should be noted that while Johannes also showed prior to this explanation an algebraic description of a vector in R^4, he did not need this representation to make his abstract generalization. Nevertheless, an algebraic description is needed to produce a specific set of linearly independent vectors.

The University students also struggled when encountering a more novel space and asked to generate examples not in R^n, but there was a distinct difference in how the students reacted to the challenge of a more novel environment. Unlike the Gymnasium students who immediately dismissed the novel space that created the difficulty, four of the five of the University students exhibited a persistent disposition to understanding and working in this new space. Consider the initial response of one University student, Martin: “I don’t know how to work with this exercises.” When asked why, Martin paused then suggested, “I can, though, take the field [pointing to the R in the prompt] to be C.” Though Martin initially hesitated in working with the exercises, he did not immediately dismiss it due to his lack of understanding. Rather, he took the time to think of what caused his difficulty and used this to persist until he had found a way to expand his thinking out of R^n. Another student Claudia had an initial reaction similar to Martin’s, then acknowledged that changing the field from the real numbers to the complex numbers would permit her to generate examples. In Figure 7, we see the example Claudia produced for a set of linearly dependent vectors not in R^n.

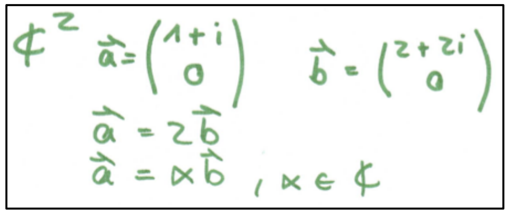


Figure 7. Claudia, a University student, persists to generate an example of a set of linearly dependent vectors in R^n.

Only one University student, Michael, failed to follow this pattern of initial hesitation before persisting to finding a solution. In the case of Michael, he did not experience the same struggle with understanding vectors not in R^n. As we saw earlier in the results, his understanding of vectors stated at the beginning of the interview included an abstract mode of description, seeing vectors as elements of a vector space satisfying a set of axioms. He used this abstract understanding to generate examples neither in R^n or C^n, but rather in the vector space of polynomials, as seen in Figure 8.

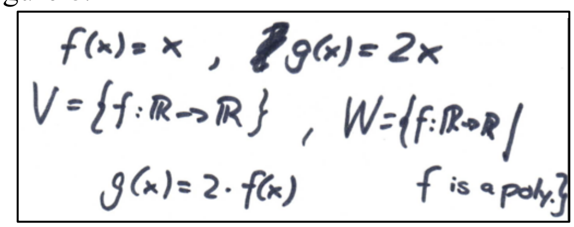


Figure 8. Michael, a University student, generates an example of a set of linearly dependent vectors not in R^n.

Thus, all University students had a disposition that permitted them to generate examples that were novel for them.

Institutional Environments and Resulting Differences

The difference in how Gymnasium and University students reacted to the challenge of extending their understanding to a novel space leads to the question, what accounts for the difference in disposition between the groups of students? What caused the University students to persevere over their hesitancy in generating examples not in R^n rather than immediately dismissing the challenge as with half the Gymnasium students in generating examples in R^4? I conjecture this is less a difference in the different content learned by these groups of student or the mastery in working with difference modes of description, but aspects of the different institutional environments that encourage this sort of tenacity.

To examine this interplay, as follows are some of my observations of the institutional norms, classroom social norms, and sociomathematical norms at both the Gymnasium and the University. As a cautionary note, this analysis is made up of generalizations that came out of observation of the classroom, field notes from these observations, informal conversations with instructors and students, and observations and conversations with student at the University in the "Math Room", a place in which students come to work on their weekly exercises and occasionally receive assistance from a tutor, often another student farther along in his or her mathematics degree. As these observations were not conducted with a rigorous analysis of the classroom and institutional norms in mind, a further, more in-depth study would be necessary to better understand how the classroom and institutional norms and individual student thinking as detailed earlier interrelate.

At the Gymnasium, students learn in an environment very similar to a traditional American high school classroom. The mathematics class observed for this study consisted of approximately 25 students. Class time was often divided between reviewing and correcting homework as a class, lecturing by the teacher, and students working on practice problems, homework, or example exercises for the Abitur exam. Lectures were primarily driven by the teacher introducing ideas, definitions, and examples on the board, but also allowed many opportunities for students to contribute questions, ideas, and summaries. For example, prior to introducing vectors, the teacher asked students when they had heard of vectors before. Two students suggested force and velocity from their physics course. Then, after the teacher presented the definition of a free vector, the teacher asked the students what might make two vectors the same. On student volunteered a group of vectors is the same if "they are all parallel, all the same direction, and all the same length" (translated). The teacher then defined equivalent vectors.

Time was also allotted in each class for students to work on exercises. These exercises were often very similar to examples provided by the teacher in the course of the lecture, as were homework exercises. Students often worked alone and in pairs on these exercises, discussing with their peers if they were confused or wanted to verify their technique and answer. Additional exercises from a textbook were frequently assigned to students as homework, which they often worked individually outside of school.

The University linear algebra course was a drastically different environment. The class was roughly 150 students. Class time was used for lectures in which the professor presented the theory of linear algebra in a rigorous, abstract format including definitions, theorems and proofs. Examples not of this abstract form were only occasionally provided. In many German University mathematics courses, students often do not have a textbook, but may have access to the professor's lecture notes in varying degrees of completion. For the class I observed, only a short summary of some highlights of the lecture notes were provided to the students through an online forum. As such, students in the lecture focused primarily on writing a complete set of notes for later work. Though the professor tried to be open to questions in the lecture, questions often went unnoticed or minimally discussed.

Student exercises worked outside of the lecture played a central role in the course. The teaching assistant and professor prepared weekly exercises sheets for the students that asked a variety of questions that extended or applied the theory presented in class. Students often worked on with peers, particularly at the Math Room. When speaking with professors and students, both said that time spent working on the exercises sheets was where the students were expected to "really learned linear algebra".

Figures 9 summarizes some of these observations, the ways modes of description were addressed in- and outside-of class and the impact on individual student thinking and disposition seen both through the institutional observation and analysis of the individual student problem-solving interviews.

University

Observations - Institution:

- Focus on algebraic and abstract modes of description – abstract favored in the lecture, algebraic and abstract in student exercises
- Exercises often extended in-lecture work
- Expectation that a majority of learning would occur in working novel exercises outside of class

Interviews - Individual Thinking and Disposition:

- Work with geometric, algebraic, and abstract modes of description
- Work with examples and exercises both similar to and different from those seen in-lecture and in exercises
- Open to new mathematical spaces and situations and exhibit a tenacious disposition

Gymnasium

Observations - Institution:

- Focus on geometric and algebraic modes of description in class and student exercises
- Exercises often repetitious of teacher-given examples and/or exercises worked in class
- Emphasis on mastery of specific types of exercises for Abitur final exam
- Expectation that a majority of learning would occur in class

Interviews - Individual Thinking and Disposition:

- Favor geometric and algebraic modes of description, rarely extended to abstract mode of description in-class and student exercises
- Work with examples and exercises that mimic those seen in-class and homework, rarely extending the exercises to novel spaces and situations
- Not as open to new mathematical spaces and situations and exhibit a less tenacious disposition

Figure 9. Summary of University and Gymnasium observations and related aspects from the analysis of student interviews

In comparing these summaries of the Gymnasium and University, we see that University students had an expectation to work novel problems on a regular basis. As such, it was part of the norms established in their linear algebra course to persist in these novel situations. In contrast Gymnasium students rarely were exposed to novel exercises, and were neither expected – by the institution or in classroom-established norms – to persist in these situations. These norms impacted students' expectations and disposition, and as such influences the students' individual thinking and activities.

Summary and Discussion

This report uses linear algebra as taught at the Gymnasium and University as a context to explore the transition between secondary school and university-level mathematics. In the course of analyzing students' modes of representation of vectors and sets of linearly independent and dependent vectors, distinct differences in the dispositions of students occurred when challenged with novel mathematical situations. While University students had the expectation of persist through these challenged established in their institutional and classroom norms, the Gymnasium students did not experience the same expectation.

One hopeful result of this study is to bring a heightened awareness of the way in which the norms established by the institution and classroom environments impact students' individual thinking. In preparing students for the transition to university-level mathematics, we must not only address the content, but the attitudes, dispositions and norms that must transition as well. By challenging students with novel mathematical situations and the expectation that they can overcome such challenges, we empower students with the disposition and expectation for themselves to persist and expand their mathematical thinking.

This research only suggests a beginning to how to measure aspects of a students' mathematical experience such as disposition and how these influence a students' mathematical work. While there has been heighted awareness to things such as disposition in K-8 education, this field of research remains relatively undeveloped at the university level. In an assessment of K-8 mathematics by the National Research Council (2001), researchers identified five strands of mathematical proficiency needed for successful mathematic learning, as depicted in Figure 10.

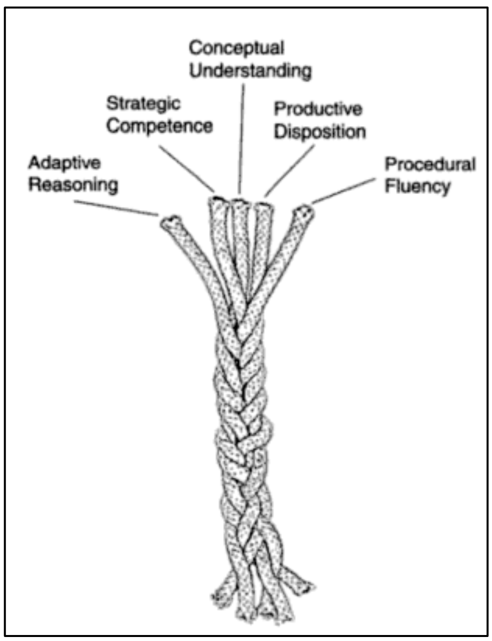


Figure 10. Intertwined strands of mathematical proficiency

This view of mathematical proficiency identifies a "productive disposition" as a key component to learning mathematics. Future research should work to expand the understanding of views such as these and what meaning they may have in undergraduate mathematics education.

References

Bernard, H. R. (1988). *Research methods in cultural anthropology*. London: Sage.

Bogomolny, M. (2007). Raising students' understanding: Linear algebra. In J. H. Woo, H. C. Lew, K. S. Park, & D. Y. Seo (Eds.), *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education, Vol. 2* (pp. 65-72). Seoul: PME.

Cobb, P. & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. Educational Psychologist, 31 (3/4), 175-190.

Dahlberg, R., & Housman, D. (1997). Facilitating learning events through example generation. Educational Studies in Mathematics, 33 (pp. 283-299).

De Guzmán, M., Hodgson, B.R., Robert, A., & Villani, V. (1998). Difficulties in the passage from secondary to tertiary education. In G. Fischer and U. Rehmann (Eds.), *Proceedings of the International Congress of Mathematicians, Volume III*. (pp. 747 - 762). Berlin.

Dorier, J.-L. (2000). *On the teaching of linear algebra*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hillel, J. (2000). Modes of description and the problem of representation in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra*, pp. 191-207). Dordrecht, Netherlands: Kluwer Academic Publishers.

Kilpatrick, Jeremy; Swafford, Jane; Findell, Bradford (Eds.); Mathematics Learning Study Committee, National Research Council (2001). *Adding It Up: Helping Children Learn Mathematics*. Washington, D.C.: The National Academies Press.

Mason, J. & Watson, A. (2008). Mathematics as a constructive activity: Exploring dimensions of possible variation. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp. 245-252). Washington, DC: The Mathematics Association of America.

Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, Documentation and Education Information Service. (2012). *Basic structure of the education system in the Federal Republic of Germany*. Bonn, Germany.

Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra*, pp. 191-207). Dordrecht, Netherlands: Kluwer Academic Publishers.

Tall, D. (1991). *Advanced mathematical thinking*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Weber, K., Porter, M., & Housman, D. (2008). Worked examples and concept example usage in understanding mathematical concepts and proofs. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp. 245-252). Washington, DC: The Mathematics Association of America.

A FIRST LOOK AT HOW MATHEMATICIANS READ MATHEMATICS FOR UNDERSTANDING

Mary D. Shepherd
Northwest Missouri State University

As students progress through the college mathematics curriculum, enter graduate school and eventually become practicing mathematicians, reading mathematics textbooks and journal articles appears to comes easier and these readers appear to gain quite a bit from reading mathematics. This preliminary study was designed to help us begin to understand how more advanced readers of mathematics read for understanding. Three faculty members and three graduate students participated in this study and read from a first year graduate textbook in an area of mathematics unfamiliar to each of them. The reading methods of the faculty level mathematicians were all quite similar and were markedly different from all the students the researcher has encountered so far, including the more advanced students in this study. A proposed Mathematics Reading Framework is given based on this study and years of observations of first-year undergraduate students reading their mathematics textbooks.

Key Words: Expert readers, tertiary reading, tertiary mathematics

Introduction

Many would agree that reading is critical for gaining understanding within a discipline. Yet, most teachers of first-year college level mathematics courses are well aware that even if they ask or require their students to read from their textbooks, few students do so with understanding. Students complain about how hard it is to read their mathematics textbooks, and it appears that even good readers in general do not read their mathematics textbooks well (Shepherd, Selden & Selden, in press). But as students continue in mathematics courses through undergraduate and graduate work, and eventually become mathematicians, somehow they "learn" to read mathematics textbooks and similar writings in journals with deep understanding and even enjoyment.

Only a little research seems to have been done on how students read their mathematics textbooks. Besides the research of Shepherd, et al. (in press) mentioned above, Osterholm (2008) surveyed 199 articles having to do with the reading of word problems, but found little about reading comprehension of more general mathematical text. He has done several studies on secondary and university students' reading of mathematical text (Osterholm, 2005, 2008) using passages written especially for that research, but not using actual textbook passages.

There has also been an interest in, and some research on, how students read their science textbooks. The journal *Science* had a special section devoted to research on, and to the challenges of, reading the academic language of science. It was noted that, while students have mastered the reading of various kinds of English texts (mostly narratives), this does not suffice for science texts that are precise and concise, avoid redundancy, use sophisticated words and complex grammatical constructions, and have a high density of information-bearing words (Snow, 2010, p. 450), much of which is common to mathematical texts.

Weinberg and Weisner (2010) have introduced a framework for examining students' reading of their mathematics textbooks. A major part of their perspective is an emphasis on the richness of personal meanings that readers construct, as opposed to the proximity of those meanings to the author's meaning or the meaning in the text (as interpreted by the mathematical community).

Is there some "thing" or combination of things that mathematicians "do" as they read that helps them understand better? Maybe mathematicians are better at monitoring their own personal understanding and have confidence that they can "fix" any misunderstanding. The questions that motivates this study: (1) Are there obvious differences in the reading strategies of mathematicians versus first year undergraduate students, and (2) If there are differences, which differences appear to be significant when the purpose of reading mathematical text is to learn from it?

Literature & Theoretical Perspective

Reading involves both decoding and comprehension. On the comprehension side of the coin, research has identified several strategies that good readers employ as they engage with text (Flood & Lapp, 1990; Palincsar & Brown, 1984; Pressley & Afflerbach, 1995). These strategies depend on the individual reader, the reader's goals and the material being read.

McCrudden, Magliano, and Schraw (2011, p. 2) conceptualize reading "as a goal-directed activity in which the reader uses text to accomplish some task and contend that successful reading comprehension is contingent upon a reader's ability to identify *text relevance*," where text relevance refers to "the instrumental value of text information for enabling a reader to meet a reading goal."

The theoretical perspective used herein is aligned with the view that reading is an active process of meaning-making in which knowledge of language and the world are used to construct and negotiate interpretations of texts (Flood & Lapp, 1990; Palincsar & Brown, 1984; Rosenblatt, 1994). Yet, it appears that for many first year undergraduate students, a major factor in their ineffective reading is a lack of sensitivity to their own confusion and errors and/or an inappropriate response to them (Shepherd, Selden & Selden, in press).

Research Questions

There is considerable reason to believe that most mathematicians can read mathematics textbooks and other mathematical writing effectively. This must be done, not only to teach new courses, but to support a mathematician's mathematical research. However few mathematicians seem to have received any instruction in reading mathematics and most seem to have tacitly learned effective reading. Although we would like to eventually know why mathematicians appear to be effective readers and first year college students are not, we limit our research question for this preliminary study to attempting to understand some differences that mathematicians have in approaches to reading mathematics versus both first-year and advanced mathematics students and whether any observed differences seem to contribute to mathematicians' apparent ability to learn from reading mathematical text. Our questions are, then:

(1) What are the obvious differences in the reading strategies of mathematicians versus first year undergraduate students (if any), and
(2) If there are differences, which differences appear to be significant in reading mathematical text for the purpose of learning from it?

Research Methods

The participants were three graduate level students and three faculty members at a large southwestern university. Initially four students and four faculty members participated, but the recordings of one of each type were not usable. One student was a masters level mathematics student, the other two were both pursuing PhD level work in mathematics education. The three faculty members were all experienced teachers and researchers. None of the participants was female.

Each participant attended a single 2 hour interview/reading session conducted by the author. After completing in initial questionnaire, all participants were asked to read (and think) aloud from *Lectures on Differential Geometry* (Chern, Chen & Lam, 2000) starting at the beginning of the book and were given instructions that they were to read to learn the material. The scenario given each of them to motivate this need to learn was that he had been chosen to teach a differential geometry course and this was one of the books he had chosen from which to begin to learn the material. Two of the faculty members had taken coursework in Differential Geometry many years ago (24 and 36), but none had done research in the area.

The students reading sessions were done first as a pilot study and from them the author had available to the mathematicians definitions that might need to be reviewed but were not in the textbook chosen. All the reading/interview sessions were video recorded and initial questionnaires were given to assess background and teaching/research experience of each participant. The reading portion of each session lasted 50-60 minutes. At the end of each reading session, the faculty members were asked to create a homework set over the material they had read. Then a few debriefing questions were posed to them.

All written materials were collected (notes and scratch work, questionnaires and homework sets). The reading portion of the video recordings was transcribed and initial coding of observations was made while transcribing. Since the participants were reading from a book, the "book" text was placed in a column on the left and what the participant said (when not actually reading) was placed on the right. Portions of the text that were skipped were coded in blue typeface, portions of the text that were "read with meaning" (explained below) were coded in red typeface. Rereading of text was coded with yellow highlighting on the right side. Green indicates how notation was read when not a standard reading. See Figure 1 below.

The number □□ is called the i-th coordinate of the point □∈□□.	…x sub i…
	Okay
For any □,□∈□□, □∈□, let (□+□)□=□□+□□, (□□)□=□□□.	For any x, y we make it … obviously a linear space…clear
	So
This defines addition and scalar multiplication in □□, making □□ an m-dimensional vector space over □.	Okay

Figure 1: Sample coding of reading transcript from Reader 6.

After the initial transcribing, another pass was made through the transcripts to time pauses and to begin to code the types of "think aloud" comments made by the readers. The results related to the comments are not presented here.

Observations and Results

The author has been observing first year undergraduate students read for approximately 10 years. One set of these observations was analyzed in Shepherd et al (in press). Many of the students observed over the past 10 years will essentially read straight through the textbook, often not looking at figures or graphs, rarely stopping to work examples or to try to make sense of formulas or definitions given. Most have little trouble decoding the symbols present, and when given help if needed, they quickly pick up on the decoding. Many will attempt paraphrases, often incorrectly, and a few will recall something (superficially) similar from their prior experiences. Few go back in the reading more than one or two sentences if something does not make sense or an apparent confusion surfaces. Most claim they have never really read their mathematics textbook with any understanding using the text mostly to find examples similar to assigned homework problems.

The graduate mathematics students in this study used some techniques and strategies similar to the first-year undergraduate students, although they were much more sensitive in monitoring their own comprehension. These graduate students essentially read the material still word for word as undergraduate students appear to do and read each symbol correctly. But unlike first year undergraduate students, the graduate students rarely read more than one or two sentences before stopping to think, reread, or work on some concept not understood. One of the graduate students never went more than 3 lines without some statement, reflection or comment on the reading. The graduate students also worked through the problems or examples on their own while undergraduate students appear to do only when encouraged to do. They were willing to spend long periods making sense of the reading or notation. One student spent 16 minutes understanding some notation and a figure. These advanced students could express awareness of incomplete understanding but make a judgment to read further, keeping in mind the less than full understanding.

In contrast, the mathematicians rarely read word for word over a long passage. They could quickly skim familiar passages and summarize efficiently what was skimmed. They frequently read "meanings" instead of the words or symbols that appeared on the page. This occurred many times with each of the expert readers. The mathematicians stopped frequently to think, reread or work on a concept or notation not understood. The average number of lines (as measured on the transcript) read without pausing for each of the six readers was under 3 lines. The mathematicians were also quite willing to spend long periods of time, particularly when reading/understanding examples. The mathematicians were also willing to suspend their need for complete understanding to read further.

Reading for meaning

The most obvious difference between the two types of readers in this study and the first year undergraduates of previous studies was the ability demonstrated by the mathematicians to "read the meaning" instead of just the symbols in a passage. Each of the three mathematicians "read the meaning" several times. Sometimes the set of symbols read with meaning was complex (see Figure 4), sometimes it was just something like *(□, □□)* which was frequently read as "the coordinate chart” or with a pronoun reference (such as “it” or “they”). Figures 2, 3 and 4 give some examples of this “reading the meaning.”

define the coordinates of *y* to be the coordinates of □=□□(□)∈□□, i.e. □□=(□□(□))□, □=1,⋯,□. The □□, □=1,⋯,□, are called the **local coordinates** of the point □∈□.	A point in phi sub U of y (*with no pausing*) So the ith coordinate is the ith coordinate of the image. So. (*1 second pause*)…Alright, so for u, each u you have local coordinates.

Figure 2: Example of Reader 5 reading meanings instead of the actual symbols.

Since □□°□□−1 and □□°□□−1 are homeomorphisms inverse to each other, □□ and □□ are continuous functions, and □□□1□1,⋯,□□,⋯,□□□1,⋯,□□=□□,□□□1□1,⋯,□□,⋯,□□□1,⋯,□□=□□.	Those (*no pausing*) If we compose in the proper way, so we will get projection to the corresponding coordinate either way.

Figure 3: Reader 6 sample of reading the meaning. The actual vocalizations are in black.

for □,□∈□□+1−0, □~□ if and only if there exists a real number a such that □ = □□. Obviously, ~ is an equivalence relation. For □∈□□+1−0, denote the equivalence class of x by □= □1, . . . , □□+1. The *m*-dimensional projective space is the quotient space □□=□□+1−0 / ~ =□□∈□□+1−0	Two points are equivalent if and only if one is uniquely a multiple of the other. That’s an equivalence relation. And then the projective space is the quotient space by this equivalence relation.

Figure 4: Reader 4 sample of reading with meaning.

In most cases, there was little or no pausing (less than one second) from apparent "seeing" of the notation to the "reading with meaning". This might indicate that somehow this translation of symbols to meaning is nearly automatic for mathematicians.

Skimming

Passages of familiar information were often skimmed and quickly summarized by the mathematicians. One mathematician skimmed nearly three transcripted pages of initial material about vector spaces and metric spaces, giving an outline summary as he went. All three mathematicians read no symbols in the description of the metric and made comments similar to the one noted in Figure 5.

For □,□∈□□, define □□,□=□=*1*□*(*□□−□□*)2* It is easy to verify that the function □*(*□,□*)* satisfies the following three conditions: 1) □*(*□,□*)≥0*, the equality holds if and only if □ = □; 2) □*(*□,□*)*=□*(*□,□*)* 3) for any □,□,□∈□□, we have the inequality □*(*□,□*)*+□*(*□,□*)*≥□*(*□,□*)*	And then it has a distance metric, so it's a metric space. And there are the properties of a metric space.

Figure 5: Reader 5 .

In contrast, the graduate students did not appear to skim at all, although this may be because they did not appear to be as familiar with some of the basic material such as the defining characteristics of a metric space, and did not skim these lines. They did not necessarily read each property in its entirety with the symbols (see Figure 6) but did try to give some meaning to complex symbols. All the graduate students commented on having seen some of this material before, in either a linear algebra or analysis class. Figure 6 shows how one graduate student read the definition of the metric. Others commented on recognizing the triangle inequality, but then went ahead and read the symbols.

Besides this linear structure, □□ also has a standard topological structure. For □,□∈□□, define □□,□=□=*1*□*(*□□−□□*)2*	to be the metric? And, uhh, this defines the norm. d of x y equals the sum of all the, um, differences between each of the ith coordinates…squares…squares of the differences of each of the ith coordinates.

Figure 6: Reader 1 (graduate student) reading the meaning of complex symbols.

Willingness to spend time to understand

Both graduate students and mathematicians were willing to engage for long periods with parts of the material to help their understanding. All the graduate students spent several minutes trying to understand the following passage. One spent 16 minutes, the other two only slightly less. There was a figure in the text also that the graduate students used extensively in making sense of the following passage.

> Suppose *(□,□□)* and *(□,□□)* are two coordinate charts of M. If □∩□≠∅, then □□*(□∩□)* and □□*(□∩□)* are two nonempty open sets in □□, and the map
>
> □□°□□*−1*|□□*(□∩□)*:□□*(□∩□)*→□□*(□∩□)*
>
> defines a homeomorphism between these two open sets, with inverse given by
>
> □□°□□*−1*|□□*(□∩□)*

This passage was not dwelt upon by the mathematicians. One skimmed and summarized it, the other two read with meaning parts or all of it. This was likely due to the greater experience of mathematicians reading or teaching about these types of function compositions.

The mathematicians did spend considerable time with the examples, though, which the graduate students did not get to when they read. Three complete examples were given in the reading, one was trivial—Euclidean m-space and the mathematicians read it and went on. The other two, the m-dimensional sphere and the m-dimensional projective space did engage the mathematicians. Figure 7 is the transcript for Reader 4 as he engages with the unit sphere. He was more likely to skim and summarize than the other two mathematicians, and his pauses were longer.

Example 2. Consider the m-dimensional unit sphere □□*={*□∈□□*+1*|□*12+* ⋯*+*□□*+12=1}* For m = 1 take the following four coordinate charts: □*1*□∈□*1*□*2>0,*□□*1*□*=*□*1*□*2*□∈□*1*□*2<0,*□□*2*□*=*□*1*□*1*□∈□*1*□*1>0,*□□*1*□*=*□*2*□*2*□∈□*1*□*2<0,*□□*2*□*=*□*2* (Figure 2 is here which is the unit circle with U1 and V2 shown)	In the m dimensional unit sphere, (*9 second pause)* We consider the following four coordinate charts. And the charts when m is one (drawing on scratch paper) (*6 second pause*) U1, points…(3 seconds) that avoid x = 0. (1 second)

Obviously {□*1*, □*2*, □*1*, □*1*} is an open covering of □*1*. In the intersection □*1*∩□*2*, we have (see Figure 2) *□2=1−(□1)2>0□1=−1−(□2)2<0* These are both □∞ functions thus (□*1*,□□*1*) and (□*2*,□□*2*) are □∞-compatible. Similarly, any other pair of the given coordinate charts are □∞ -compatible.	And the coordinate chart is…(appears to be thinking about what these charts are—12 seconds) (writing and mumbling) x 1 squared plus x 2…squared… …(??) one,. okay and we have phi okay U 1 (writing) x1 with xsquared > 0. Phi of x equals x1…U2 with x squared < 0…still x 1. 3 the V1 x in this one with x1 > 0, and the map is x squared. And V2 x in S1 with x1 < 0 and the map is still x squared (pause 15 seconds—confused). So here I have a problem with notation (*pause 10 seconds*, flipping back 2 pages, then

	back to place) We need to clarify x being a 2-dimensional point, what's the meaning of x squared. (mumbling as writing) This is d??? x1 squared plus x2 squared….That can't be. That would be the natural assumption, but this is equal to 1 on S1. (looking at figure 2). So let's see if this makes it clear. (pointing to figure 2. *46 seconds of thinking*) So the relation is x…x1 squared..Oh, that's just x2, not x squared…I'm getting all mixed up on notation I have never used. Okay. So this is completely clear (crosses of on

	notes)…So it's either the first coordinate or the second coordinate. U1 has x2 positive, so that's the upper half…lower half ….(adding these to notes) right half..and left half. …its… Because all four of them cover it (circling pencil). (*pause 2 seconds*) And then the relations are C infinity. (*2 seconds*) So we say these are C-infinity compatible. That's okay.

Figure 7: Reader 4 script for Example 2.

One expert reader was constantly searching for examples. After reading the definition of m-dimensional projective space he envisioned the 2-dimensional projective plane and used it to help enrich his understanding of the coordinate charts given. The sample in Figure 8 is only his initial look at the lower dimensional case. He continues to use this for several more minutes while reading about the coordinate charts in projective space.

The m-dimensional projective space is the	

quotient space □□=□□+1−0 / ~ =□□∈□□+1−0	The quotient space of the R m+1 minus 0 over this relation. Uhhh. (*4 seconds*) Uh huh. Yes, okay, so for example. …He doesn't give any examples. Well I'm trying to make it for m = 2 and m+1 =3. So then we are talking about 3 dimensional space minus the origin and the two the two points or two vectors in it are equivalent if they are along the same ray, so positive or negative, they are still. Yeah, if it were…if they were equivalent only if this number [a] is positive it would have been just a sphere. But now they are equivalent for any a, so the opposite points are still equivalent, so it's like we take a sphere and and and and make and glue together each two opposite points. That will be the projective plane…projective plane I guess. Yes, okay. And okay, so then we go with projective space of dimension m….the m dimensional projective space.

Figure 8: Reader 6 as he looks at a lower dimension case to help his understanding.

In addition readers in both groups were very cautious about their own understanding and frequently adjusted their interpretation to match more closely that of the authors of the textbook. For example, the notation in the passage read was "non-standard" in the use of superscripts in place of subscripts. The following definition was given in the text: Let □□={□=(□1,…,□□)|□□∈□, 1≤□≤□} The superscripts caused some comments and minor confusion but the notation was figured out by all the readers eventually. An example of this working out of confusion also appears in the script in Figure 7 above.

The types of passages that the graduate students and the mathematicians spent time on to understand were different and this appeared to be based on experience. Graduate students spent time understanding notation, experts spent time understanding the examples and the "calculations" needed. Graduate students read with only slight paraphrasing/summarizing the "should be known" sections that the experts quickly skimmed and summarized.

Suspension of Need for Complete Understanding

Both sets of readers in this study were diligent about monitoring their personal understanding of the material. And each set made judgments about whether to work further when their level of understanding did not meet their personal expectations.

> **Reader 2 (Graduate Student):** So, as I'm reading this they are defining a homeomorphism between two open sets. Although I don't completely understand, I'm going to see later on where it's used and how it's useful. And if I need to come back to this, then I will. There's a picture here that might clarify things.

Reader 3 (Graduate Student): Suppose r is a real valued function defined on an open set U subset of Rm. Thinking. What does that actually do for us. Well, I mean what does it do for us, okay, so defined on open (pause). Okay, what I'm going to do is read ahead and see if I get anything out of that, if not I'm going to backtrack.

Reader4 (Faculty): So the question is what do we mean by ..by…what is u j for j less than zero? u is a point in R4 minus 0 so it's (flips back a page) meanless (flips back another page then to current page). Okay, I'm going to have to ask about this example. (Writes a note on scratch paper). It bothers me, but I'm not going to spend the whole day with it.

One of the mathematicians, while recalling some related fact realized it might not be relevant and was able to "leave it alone."

Reader 6: Like all top…ehh..countable topological basis, for example m dimensional space has a countable topological basis. Even infinitely dimensional, like, uhh, uhh, like Hilbert space probably has, but it may not but then (??) I remember it from my studies of..in functional analysis, if you don't assume that, some things become extremely difficult. So, yeah, we better just leave this case alone.

Finally, our expert readers and graduate students, when not suspending their need for understanding were very willing to use external sources to refresh/review unfamiliar or forgotten concepts. Traditional books and Google were mentioned as sources for finding quick summaries/definitions of unfamiliar or forgotten concepts.

Towards a conceptual framework for reading mathematical text.

The theoretical perspective that has been used in approaching how someone reads for the purpose of learning mathematics is aligned with the view that reading is an active process of meaning-making. We can now see some trends in how various readers approach mathematical text. There is an element of reading fluency, reading orally with speed, accuracy, and proper expression[1], that was exhibited by all readers. But more advanced readers exhibit an ability to read the meaning, not just the symbols that novice readers of mathematics do not. It also appears that advanced readers monitor comprehension better and have learned that they can often "fix" a comprehension failure. Table 1 below gives a comparison of the observations of the different "levels" of reading mathematics and the mathematical level of the readers in each category. The "codes" on the left refer to the dimension in the Mathematics Reading Framework described below.

From this categorization we might suggest that there is a multidimensional continuum of observed reading behaviors for reading mathematics. We emphasize three of these dimensions, mathematical fluency, comprehension monitoring, and engagement with non-text material such as tables, graphs, worked examples or exercises.

Novice Readers First Year Undergraduate Students	Intermediate Readers Early Program Grad Students	Expert Readers Faculty (n=3)

[1] National Institute of Child Health and Human Development. (2000). Fluency.

		(n=3)	
MF	Nearly always read word for word sometimes stopping to ask "How do you say that?"	Nearly always read word for word.	Read small sections word for word, but often read the "meaning" and not the words/symbols.
E	Rarely stop to look at graphs or figures	Stopped to understand figures within the context of what was being read	Skimmed figures to make sure see that they matched understanding
CM	Often read long passages without stopping to check understanding	Read only one or two sentences before checking understanding.	Might read as much as a paragraph, but had no compunction about stopping to check understanding.
CM	Rereading usually only back 1 or two sentences	Might reread a sentence 4 or 5 times while trying to make sense of it	Might reread a sentence 2 times to check understanding.
MF	Paraphrasing occurs, often missing some conditions.	Not much paraphrasing observed	Passages on "known" material often quickly skimmed/summarized but not read directly
E	Not much engagement with trying to understand the new material	Long passages of engagement with understanding symbols/meaning	Long passages of engagement with understanding symbols/meaning.
CM	Might ask about unknown terms/symbols, but would not usually go to index to find out more.	Willing to look up unknown terms	Willing to look up unknown terms
E	Sometimes will use scratch paper.	Willing to use scratch paper freely for exploration	Willing to use scratch paper freely for exploration
CM	Often apparently not aware of lack of understanding	Aware of lack of full understanding but willing to read further.	Aware of lack of full understanding but willing to read further.
CM	May recall something (superficially) similar but often unsure how to relate to current reading.	May recall from a previous class, and willing to look up again.	May recall and can decide whether relevant or not.
E	Not willing to create example in general	(no observations here)	Willing to create or extend an example beyond what is given in the text.

Table 1: Observed differences in reading strategies in different groups of readers.

Mathematical fluency (MF)[2]. On the novice end of the spectrum this might include reading the words and/or symbols haltingly or asking how to "say" that. Intermediate readers would read the words and symbols smoothly. Expert readers would read the meaning in place of some words and symbols.

Comprehension monitoring (CM). Novices are often unaware there is a comprehension issue, or may have an overly optimistic view of their comprehension. Intermediate readers are aware of comprehension failures, and may be willing to fix them, but are sometimes unsure how to fix the problem. Expert readers are very aware when there is a comprehension failure and are willing and confident that the problem can be fixed.

Engagement with non-text material (E). Novice readers only minimally engage with figures or tables and will "read" an example but are not likely to work an example on their own. Intermediate readers willingly engage with tables and graphs to enhance their comprehension. Expert readers are likely to skim tables and figures to confirm/check their understanding and willingly engage with examples to enhance their understanding.

Table 2 below is a summary of these three dimensions a proposed Mathematics Reading Framework.

Mathematical Fluency	Comprehension Monitoring	Engagement with non-text material
• Novice: reading words/symbols haltingly	• Unaware of comprehension issues	• Minimal engagement with tables/graphs, will not usually work through worked examples.
• Intermediate: Reads words and symbols smoothly	• Moderately aware of comprehension failure and willing to fix but sometimes unsure how.	• Willingly engages with tables and graphs to enhance comprehension
• Expert: Reads the meaning frequently instead of words/ symbols.	• Fully aware of comprehension failure and confident in ability to fix if needed.	• Skims graphs/tables to check comprehension, willing the spend long time on examples to enhance comprehension.

Table 2: Proposed Mathematics Reading Framework

Summary and answers to research questions

So, where does this leave us in regards to our research questions? (1) What are the obvious differences in the reading strategies of mathematicians versus first year undergraduate students (if any), and (2) If there are differences, which differences appear to be significant in reading mathematical text for the purpose of learning from it? There do appear to be obvious difference in reading strategies. It would be presumptuous to indicate the significance of the difference in

[2] The codes MF, CM and E are used in Table 1 to identify examples of each dimension.

such a small study, but it appears at least three overarching differences exist. First, expert readers read the meaning, not necessarily the symbols often while reading. This was the most noticeable difference between the different readers and was consistent across the three (and even the fourth whose interview was not recorded property). Second, the expert readers clearly are very conscious of their own understanding and trying to match their understanding to that of the author. Experts appear to search out ways to "fix" any perceived misunderstandings, whether working through notation or going to outside sources. Finally, the readers in this study were willing to spend large amounts of time working through notation and examples. The graduate students spent more time working to understand notation, the mathematicians spent more time with the examples. The experience of the mathematicians may have been the source of some of the differences in the types of passages the readers engaged with for long periods of time. Or it might be the case that as readers mature in their ability to read mathematical text, they learn where to focus their attention and work for the most complete understanding.

Implications for Further Research and Teaching

This research project is a preliminary step in understanding the broad scope what it means to read mathematical text for understanding. This is an initial pilot research project to begin to understand the "expert" side of reading mathematical text. Previous research has focused on the "novice" or first-year undergraduate course student. From this study, a preliminary Mathematics Reading Framework is proposed and given in Table 2. As research into reading mathematics textbooks continues, there are opportunities to understand not only what experts "do" differently, but how they learn to do this and what steps or phases of learning to read occur between novice and expert. We can also anticipate the integration of reading for understanding with learning theories.

This current research has strong implications for teaching as we ask how can we move our students along in the Mathematics Reading Framework. There are also implications as we design tasks and textbooks both traditional and online. What can we as teachers do to help our students move toward the "expert" end of the different dimensions of the Mathematics Reading Framework? In fact, the author, who has been giving reading guides to students to help them learn to read their mathematics textbooks, has modified the types of reading guides given to students to help them progress from novice readers towards expert readers of their textbooks. They are asked to frequently stop and "do something" to check their comprehension and they are now encouraged to read the notation appearing in limits with the symbols $\to\pm\infty$ or $=\pm\infty$ with "increasing or decreasing without bound" language instead of "infinity" language. Is this effective? What learning trajectories exist for moving a student from novice reader of mathematical text to expert reader? How does reading mathematics for understanding integrate with learning theories.

The text chosen for this study was one on a topic unfamiliar to the readers. There were no theorems in the portion read. Does familiarity with the topic change the reading strategies/actions? It was clear that both the graduate students and the mathematicians were excited when they finally understood some concept they had struggled with. If one of the reasons mathematicians read more effectively is because they have had positive reinforcement that they can learn from reading, how do we achieve similar positive reinforcement with lower level students?

References

Chern, S., Chen, W., & Lam, K. (2000). *Lectures on differential geometry.* World Scientific.

Flood, J., & Lapp, D. (1990). Reading comprehension instruction for at-risk students: Researchbased practices that can make a difference. *Journal of Reading, 33,* 490-496.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2011). *Text relevance and learning from text.* Charlotte, NC: Information Age Publishing, Inc.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Retrieved March, 24, 2012, from http://www.nichd.nih.gov/publications/nrp/smallbook.htm.

Osterholm, M. (2005). Characterizing reading comprehension of mathematical texts. *Educational Studies in Mathematics, 63,* 325-346.

Osterholm, M. (2008). Do students need to learn how to use their mathematics textbooks? The case of reading comprehension. *Nordic Studies in Mathematics*,13(3), 7-27.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1,* 117-175.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum Assoc.

Rosenblatt, L. M. (1994). The transactional theory of reading and writing. In Ruddell, Ruddell, & Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 1057-1092). Newark: International Reading Association.

Shepherd, M., Selden, A., & Selden, J. (to appear). University students' reading of their first-year mathematics textbook. *Journal of Mathematical Thinking and Learning.*

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328, 450-452.

Weinberg, A., & Wiesner, E. (2010). Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics*, 76, 46-63.

.

EXPERTS' REIFICATION OF COMPLEX VARIABLES CONCEPTS: THE ROLE OF METAPHOR

Hortensia Soto-Johnson, Michael Oehrtman, Kristin Noblet, Lee Roberson, & Sarah Rozner
University of Northern Colorado

Using a theoretical perspective of embodied cognition, we explored how six experts integrated metaphors to reason and communicate about arithmetic and analytic complex variables concepts. We found that experts who displayed evidence of reification of a complex variables concept or had a need to use a concept imparted their sense of understanding through enacted metaphors. These metaphors were often invented or reinterpreted, based on personal experiences and created to convey nuances of the experts' understanding to students. The experts appeared conscientious of using metaphors relevant to their own students. This research may support practitioners' efforts to create opportunities for students to create or reinterpret experts' metaphors into personally meaningful metaphors that both capture important mathematical concepts accurately and align with their own understandings, experiences, and culture. Further research may investigate how technology may serve as a tool for such an endeavor.

Keywords: Complex variables, Embodied cognition, Mathematicians, Metaphor

Introduction and Literature Review

Given the significant body of literature investigating students' understanding of real numbers, ranging from the meaning behind arithmetic operations (Sowder, 1992) through analysis of real-valued functions (Alcock & Simpson, 2004), it is natural to seek extensions of these studies to complex numbers, their operations, and functions. Such studies may provide insight into processes of generalization and abstraction as well as into potential ways to strengthen students' understanding of, representations of, and fluency with operations on complex numbers. They may also provide insight into the development of intimately related concepts in grades 9-12 and undergraduate curriculum involving functions, vectors, matrices, and transformations. Our research is designed to contribute to the literature on teaching, learning, and understanding undergraduate mathematics. This report is part of a larger exploratory study in which we investigate experts' geometric reasoning about complex variables in an effort to create a framework based on empirical evidence that describes how one perceives and reasons with central ideas from complex variables. In this paper we address the research question: What is the nature of experts' use of metaphor in conveying their perceptions about the arithmetic of complex numbers and analysis of complex valued functions?

There is limited research investigating the understanding of complex variables, but there are a handful of empirical studies that have begun to pave the road in this domain. In their work on embodied cognition, which we discuss in more detail in the following section, Lakoff and Núñez (2000) presented a framework for the conceptual development of complex numbers. Their framework blends the real number line, the Cartesian plane, and rotations with the use of metaphor for number and number operations. They began by imparting physical meaning to the product of a real number x with -1, as a rotation of 180° to obtain $-x$. Similarly they depicted multiplication by i as a clockwise rotation of 90°. Lakoff and Núñez' perception of these numbers as operators that transform an object might suggest that if students perceive multiplication by –1 as a rotation of 180°, then they might easily recognize that multiplication by i results in a clockwise rotation of 90°.

In an effort to test this conjecture, Conner et al. (2007) explored ten preservice secondary teachers' understanding of complex numbers via in-class video recordings, in-class assessments, homework assignments, and students' responses on the final exam. Contrary to Lakoff and Núñez' (2000) framework, the researchers found that the participants viewed multiplication by –1 in the complex plane as a reflection rather than a rotation. This result could be attributed to the fact that the students focused on the real number line rather than the entire complex plane. The data also implied that the students failed to illustrate multiplication of complex numbers with any geometric interpretation on the complex plane, did not recognize that if z was a solution to a quadratic equation then the equation evaluated at z would yield zero, and viewed a complex number as a pair of real numbers rather than as a single entity. On a more positive note, the prospective teachers did demonstrate how to use the complex plane to illustrate addition of complex numbers using vectors or by decomposing complex numbers into real and purely imaginary components.

In another study related to complex numbers, Danenhower (2006) examined the ability of Canadian undergraduates, enrolled in a complex variables course, to convert instantiations of the fraction $\frac{a+ib}{c+id}$ to either Cartesian ($x + iy$) or polar form ($re^{i\theta}$). The varied fraction representations included taking the modulus of the numerator, raising the factors in the numerator and denominator to a power, expressing the denominator in terms of sine and cosine, and combinations of these forms. The undergraduates worked flexibly with complex numbers when represented in Cartesian form, but this was not the case with polar representations because the participants were not comfortable with trigonometry. Students' relative comfort level with $\mathbf{R}$ and $\mathbf{R}^2$ could partially account for the students' preference of the Cartesian form over polar form. Danenhower's findings also suggested that the undergraduates did not attend to geometric representations of the complex number represented by the fraction, which could have alleviated much of the computational effort. His work suggests that his participants were limited to viewing i as a static object, and did not possess a dynamic view of multiplication by i as an operator, which acts on other objects. One of the most significant contributions of Danenhower's work was his observation of a phenomenon, which he referred to as "thinking real-doing complex." In his dissertation, Danenhower (2000) explained that this phenomenon emerged when students applied their understanding of $\mathbf{R}^2$ while working with complex valued expressions and functions. For example one student attempted to determine if a complex valued function was differentiable by simply inspecting the mapping of the function. The students attempted to use a strategy a of drawing a line as they might, although not always useful, in $\mathbf{R}^2$.

Nemirovsky et al. (in press) presented promising results, based on a teaching experiment, of how preservice secondary teachers were able to view i as a an object that causes a certain behavior in complex numbers depending on the operation. The goal of their teaching experiment was to provide students with an instructional sequence where they created conceptual meaning for adding and multiplying complex numbers. As part of the teaching experiment the participants used the floor as the complex plane and physically played a part in determining the behavior of multiplying $2+\frac{1}{2}i$ by i. Using methods from microethnography, the researchers generated detailed characterizations of students' gestures during short episodes of the teaching experiment. Upon summarizing and reflecting upon these gestures, Nemirovsky et al. concluded that perceptuo-motor activity was central in (1) conceptualizing, (2) communicating geometric representations, and (3) creating a learning environment that influenced the development of structural components behind adding and multiplying complex numbers. These results advocate that enacting a geometric interpretation of the multiplication of two complex numbers enhances students' development of viewing multiplication by i as a dynamic process.

In a study with Greek high school students, Panaoura et al. (2006) investigated students' (N = 95) ability to navigate from an algebraic representation to a geometric representation of complex-valued equations and inequalities and vice-versa. The equations and inequalities were of the form $|z - z_0| \le k,$ where z was the variable, z_0 was a fixed complex number and k was a positive real number. The researchers administered two questionnaires; in the first questionnaire students were required to provide the geometric representation of a given algebraic equation or inequality and in the second questionnaire students were asked to produce the algebraic representation of a given figure. Both questionnaires concluded with a problem-solving task similar to the items on the questionnaire. The findings indicated that students were more successful in answering the items correctly when provided with the geometric representation. On the other hand the students were not consistent in implementing these strategies on the problem-solving tasks, which suggests "a lack of flexibility in using the geometric approach effectively with different representations of complex numbers" (p. 700). These conflicting results could be due to students' inability to connect the symbolic/algebraic, geometric, and verbal representations. Instead, the students tended to compartmentalize the different representations and did not recognize the similarity between the problem solving tasks and the original items.

The literature shows that researchers have investigated high school students', undergraduates', and prospective secondary teachers' geometric understanding of complex numbers as well as their ability to move between representations. Our research extends the literature in terms of population and content; specifically we focus on experts' geometric interpretations of complex numbers and complex valued functions. Their responses may provide insight into how educators may better expose students to complex numbers and complex valued functions in an effort to help them develop facility with multiple representations.

Theoretical Perspective

Embodied cognition serves as our theoretical perspective and stems from the theory of enactivism. This theory asserts that "the individual knower is not simply an observer of the world but is bodily embedded in the world and is shaped both cognitively and as a whole physical organism by her interaction with the world" (Ernest, 2010, p. 42). For enactivists, knowing results from interactions with the world and is impacted by the knower's previous experiences including the meshing of cultural, social, and individual events, which are not viewed as separate entities. Through this lens, learning mathematics occurs when "students change their structures, and therefore their behavior, in a complex process of interaction with their environment" (Lozano, 2005, p. 25). As such, learning or perceiving concepts and action go hand-in-hand. "*What we perceive* is determined by *what we do*, … it is determined by what we are *ready* to do. … we *enact* our perceptual experience; we act it out" (Noë, 2004, p. 1).

Acting out a concept suggests that the learner has imposed a structure on the concept. Sfard (1994) coined this as reification, which she defined as a capstone of the development from operational to structural reasoning. She complemented this definition with a focus on the particular metaphor of mathematical constructs as physical objects. Thus for Sfard, reification is the creation of metaphor, which plays a central role in the philosophy behind embodied cognition.

Sfard paraphrased Lakoff and Johnson's definition of metaphor as a "mental construction, which plays a constitutive role in structuring our experience and in shaping our imagination and reasoning" (p. 46). According to Lakoff and Johnson (1980) embodied schema also known as image schemas are the mechanism for creating metaphors. They are structures of an activity by which we organize our experiences in order to create meaning. These image schemas may not be rich in detail, but they are embodied and can encompass multiple and diverse experiences. Lakoff and Núñez (2000) extended the idea of embodied metaphors to the discipline of advanced mathematics topics. They advocated that "general cognitive mechanisms used in everyday nonmathematical thought can create mathematical understanding and structure mathematical ideas" (p. 29). Furthermore, they described image schemas as the link between language, reasoning, and vision, on which we focused our data analysis.

Research Methodology

In an effort to obtain rich data, we selected a purposeful sample of six expert participants. The participants Ricardo, Anton, Mark, and Beth were selected based on our personal interactions with them or based on student comments. Upon interviewing Beth, she suggested we interview Luke and Jane with whom she collaborates on complex analysis research. All the participants are PhD mathematicians except for Mark, who is a PhD physicist. The experts participated in a 90-minute video-taped interview, where two researchers posed questions aimed to reveal the participants' physical interpretation of arithmetic and analytic concepts related to complex variables. We informed the participants that we were investigating their geometric interpretation of complex numbers and complex variable concepts. We conveyed our interest in their use of gestures, diagrams, illustrations, and facial expressions, but we did not use the word metaphor. The participants described their connections between algebraic and geometric representations of addition, multiplication, division, and exponentiation of complex numbers. They also conveyed their geometric perceptions of continuity, the Cauchy-Riemann equations, differentiation, and line integration of complex-valued functions. Appendix 1 contains the interview items. Probing was used throughout the interview for clarification purposes or in an effort to elicit ways in which our participants might incorporate geometric or visual interpretations in explaining ideas to novices such as undergraduates.

We used phenomenological methods in our analysis, which entailed careful condensation of the data. Such analysis does not "use coding, but assumes that through continued readings of the source material and through vigilance over one's presuppositions, one can … [capture] the 'essence' of an account – what is constant in a person's life across its manifold variations" (Miles & Huberman, 1994, p. 8). These methods allow one to gain an understanding of meanings and actions. Implementing these methods began with four members of the research team transcribing, time-stamping, and conducting an initial analysis documenting where and how a participant conveyed her or his perceptions using geometric methods. After this individual analysis, as a team we repeatedly viewed every interview in its entirety and read the transcripts multiple times to determine common themes among the participants' responses. It was during this time, that we noticed the experts' repeated use of metaphor, which led to a more focused research question and more detailed analysis of the data. During the next phase of the analysis, we created profiles for each participant. These profiles consisted of a title based on the metaphor that was used, quotes, parsed segments of video-frames with arrows indicating gestures (as seen in the results section), a detailed description of the participants' enactments, and our interpretations of what we believed the participant was attempting to convey. In the language of

Barker (2009), the metaphors and diagrams alone provided us with an *idea* of the experts' geometric perception but the detailed and active gestures offered us a *sense* of their perceptions, which allowed us to better *grasp* their interpretations of the concepts.

Results

Our results suggest that participants who displayed evidence of reification of complex variable concepts imparted their understanding through dynamic representations and enactments of metaphors. The metaphors were often invented or reinterpreted, based on personal experiences, and created to convey nuances of the experts' understanding to students. Most of the experts found the questions about exponentiation of complex numbers, the Cauchy-Riemann equations, and line integration of complex valued functions novel and hesitated to create meaning of these situations. On the other hand, our participants enacted similar metaphors for the arithmetic operations of complex numbers and the continuity and differentiation of complex valued functions, which is not uncommon under the enactivist learning perspective (Lozano, 2005). In this report, we describe the metaphors used including quotes and video frames from our analysis, in an attempt to paint a picture of the responses. We begin by synthesizing the metaphors and categorizing them based on similar features, which we refer to as a cluster.

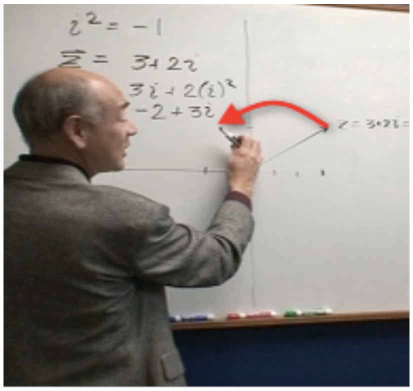
Table 1 is a summary of the clusters and examples of metaphors, which the participants used in responding to each item. The cluster "mapping" was an overarching theme in all of the items; this might not be surprising given that complex valued functions and the operation of complex numbers are often characterized as transformations of the complex plane. What might be more unexpected is that not all of the participants expressed this view consistently throughout the interview; only Ricardo provided mapping related metaphors for the Cauchy-Riemann equations and line integration items. In the remainder of this section, we provide detailed summaries of how our participants enacted their metaphors so that the reader may be able to *sense* and *grasp* the participants' geometric perceptions of these items. Due to length limitations, we only provide summaries for representative or exemplar responses.

For the arithmetic questions we provided a drawing of the Argand plane with two complex numbers z and w and asked the participants to determine where $z+w$, zw, $\frac{1}{z}$ and z^w were located on the Argand plane. We also asked them to articulate connections between algebraic and geometric representations. All the experts described addition of complex numbers in terms of vector addition and illustrated a parallelogram created by the two vectors corresponding to the two complex numbers. For example, Beth commented, "*You can think of them as vectors. So I looked at what vector z looked like.*" It is interesting that Beth no longer referred to the *complex number z*, but rather the *vector z*, which suggests she viewed complex numbers as objects. Accompanying gestures to the parallelogram model included starting at a point z and then sweeping an index finger in the horizontal direction followed by a sweep in the vertical direction to indicate adding the complex number w. Some participants used pincher fingers, formed with their thumb and index finger, or their two index fingers to denote the length from the origin to the real component of w and used their pincher fingers as a measuring tool to measure off the distance from the origin to the real component of z. Similar actions were used for the imaginary component of the complex numbers, which allowed the experts to communicate their understanding between the algebraic and geometric representations of adding complex numbers. Using the sweeping movement, Ricardo mentioned that each vector has a "*motion*" in the horizontal and vertical direction. Thus, for Ricardo, complex numbers were mobile physical objects. Both Luke and Ricardo stressed the facility of thinking about addition in terms of rectangular form and then simply adding component wise. Luke remarked, "*...you can see in the picture that if you just look at the x-components of the two complex numbers that you get the x-component of the sum that I've drawn, and similarly for the y-components.*" The predominance of these gestures suggest that the experts not only reified the complex numbers as objects but conceived of them in very physical terms.

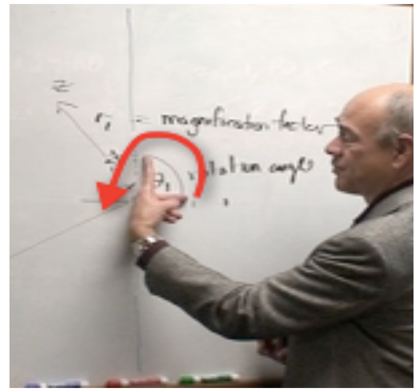
Table 1
Metaphors Used on Items

Interview Items	Cluster	Examples
Arithmetic	Representations	Cartesian, polar, points, vectors, ordered pairs, motion
	Mappings	translating, rotating, dilating, operators/operands, composing, matrices, input/output
Continuity	Discontinuity	silly puddy, tearing paper, radiation reaction and charge at rest, deforming
	Multivariable Calculus	hiking, elastic bands, parking lot
	Mappings	controlling range, pulling back with contours, sending balls to balls, archery competition, painting, close goes to close, topology, preserving coherence of plane, input/output
Cauchy Riemann Equations	Jacobian Matrix	symmetries
	Mappings	turn table, rigid body, rotating, dilating
Differentiation	Multivariable Calculus	linear approximations, slope, tangent planes, scaling, plane wave and momentum, real linear with added symmetries
	Mappings	rotating, dilating, pinwheel, bicycle wheel, turn table, clock dial, amplitwist, impacting radius and angle near a point, point by point, conformal
Line Integration	Mappings	captain's chart, ship, timepiece

Unlike the students in Danenhower's (2006) study, our expert participants recognized and mentioned the value of looking at complex multiplication and division in polar form rather than Cartesian form. In responding to the multiplication item Jane quickly said, "*think of it in polar form because that's the natural way to multiply complex numbers*" while simultaneously moving her hand in an arc motion indicating a rotation. All the participants expressed how the polar form helped one recognize the geometric action of multiplying two complex numbers and they all commented that the scaling of the vector w would depend on where z was located in terms of the unit circle. They were cognizant of the fact that if z lied inside the unit circle, was on the unit circle, or lied outside the unit circle resulted in shrinking the magnitude, not impacting the magnitude, or stretching the magnitude of w respectively. It was not uncommon for the participants to use their pincher fingers (described above) to activate the stretching or shrinking and to physically measure and keep track of the angle measures that they were adding together to obtain the resultant angle.



Frame 1: Ricardo illustrating multiplication by i



Frame 2: Ricardo rotating by the rotation angle

Figure 1. Ricardo enactments for multiplying two complex numbers.

Figure 1 illustrates two ways in which Ricardo connected and illustrated the rotation resulting from the multiplication of two complex numbers. In frame 1, Ricardo considers a specific example where $z = 3 + 2i$ is multiplied by *i*, because he wanted to share that this is the process in which he conveys multiplication to his students. He explained that he wants students to recognize that multiplication by *i* is simply a rotation of 90°. This allowed him to generalize multiplication by any complex number, which he illustrates in frame 2. From Ricardo's language of, *"... I would think of the angle for z as θ_1 and let it have length r_1, which is the magnification factor and the angle is the rotation factor, then w will spin through that much angle and at the same time multiply by that factor*" it appeared as though he viewed *z* as the operator and *w* as the operand. A follow-up interview with Ricardo confirmed this assertion. Similar to the other participants' responses, Ricardo imparted action to the symbolism involved in the polar representation i.e. magnification and rotation factor and enacted this action as he elaborated on his response. Ricardo's response to the exponentiation item has similar features using repeated composition of transformations; we invite our readers to read this in Soto-Johnson, Oehrtman, and Rozner (2011). We should note that Ricardo was the only participant to impart geometric structure on the exponentiation item.

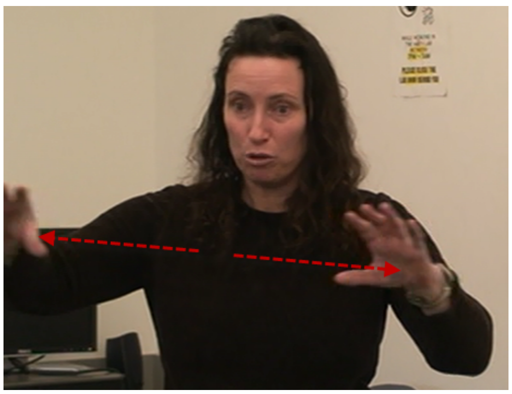
In the multiplication item, some participants discussed the rotation aspect first followed with a conversation about the dilation aspect while others reversed the order of the transformations. Similarly in the division item the experts varied the order of the reflection and the dilation. For example Mark, began by concurrently stating and writing $\frac{1}{z} = z^{-1} = \frac{1}{r}e^{-i\theta}$. He then placed his left hand with palm up and his index finger pointing at *z* above the real axis as he said, "*so there's a complex conjugation*" while flipping his palm down and fingers spread out, below the real axis. He proceeded with the comment, "*... if r_1 were greater than 1 and the unit circle were in here* (drew a circle with a radius smaller than r_1) *then there's an inversion*. *If we were inside the unit circle, we would go out* (as he motioned outward*)*." All of the participants provided similar descriptions and commented on the commutativity of the two transformations.

In the addition item, the experts clarified that using vectors is "*natural*" because students are familiar with vector addition. Since continuity of complex-valued and multivariable real-valued functions is the same, it did not seem unusual for some experts to use this concept to provide a geometric representation or explanation to convey their understanding of continuity of complex-valued functions. As with the vector notation, several participants commented that such an explanation allowed them to make connections to students' prior knowledge. Our participants also presented metaphors, which they believed would be relevant to their students' experiences. In communicating these metaphors, the experts intertwined drawings, enactments of the metaphor, and gestures. For example in her hiking metaphor, Jane explained, "… *so then if I'm thinking of the pen off the page,* [used marker positioned out from the board to trace the height values] *I might be trying to trace out on the surface, trace out those height values and see if I can draw them, I'm not taking my pen off the page now, but I, I don't want to jump my pen anywhere* [traced heights with marker, then pulled marker out from board to illustrate a jump]. *That's my analogy of- I don't want my pen to make any sudden precipitous drops. So, I often use a hiking analogy, especially in the classroom setting, because the students are familiar with contour maps and falling off cliffs or not falling off cliffs…. So if I can draw this without sudden change of altitude* [traced and pushed pen into the board], *then I'm continuous."* An interesting aspect of this 3-D metaphor was that Jane's description fit with what one would observe with two-variable functions – again because the complex-valued and multivariable real-valued functions are the same. Jane also used a parking lot metaphor, where she used her index fingers as cars on a two-lane road going in opposite directions and cars in a parking lot driving in different directions. Figure 2, depicts Jane's enactments and comments of her parking lot metaphor (the arrows illustrate the direction of her hand movement). In frame 1, Jane used her right hand to represent the car moving on the single lane-road, but it is curious that her road is not horizontal. In frames 2 and 3, Jane conveyed how big the parking space, which is evident through her hand gestures as well as her wide-opened eyes. Finally, in frame 4 Jane indicates cars traveling in different directions.

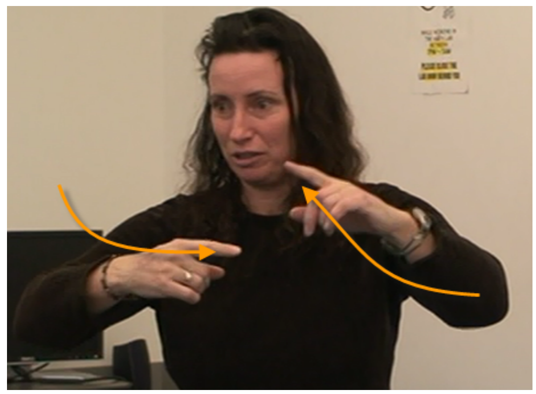
The intriguing aspect of using multivariable real-valued functions to describe continuity of complex-valued function was that the participants were thinking real while doing complex (we also witnessed this with the differentiation and integration items), which is in line with Danenhower's research (2006). They thought in terms of functions that map from $\mathbf{R}^2$ to $\mathbf{R}$ rather than from $\mathbf{R}^2$ to $\mathbf{R}^2$. We also observed this phenomenon with participants who chose to convey their understanding of continuity by discussing discontinuity though the use of metaphor. For example, Ricardo and Beth used a tearing paper and a silly putty metaphor respectively. Essentially, they both described discontinuity as objects that start close together ending up far apart. In Beth's description she commented, "*… in the analogy of not lifting up the pen is if you made a region out of silly putty, and you applied the function to every point in that region, what would that shape look like.*" As she made this statement she clasped her hands together in a horizontal position, rubbed them together as if rolling silly putty, stopped and arced her arms with hands together to indicate the mapping, then she separated her hands. She further remarked, "*Would you have to rip the silly putty to get there?*" as she put her fists together followed by pulling her hands apart in opposite directions. She completed with the statement, "*An analog to not lifting your pencil, where we usually think of discontinuity as having a break in the graph, in complex we think of there being a tear in the image*." This sequencing is illustrated in Figure 3. Beth effortlessly switched from an image mapping from $\mathbf{R}^2$ to $\mathbf{R}$ (tracing a curve on a surface) to an image mapping from $\mathbf{R}^2$ to $\mathbf{R}^2$ (separating the silly putty).

Frame 1: *"I often end up describing the difference between driving on a single lane road"*

Frame 2: *and then being on a big wide open parking space*

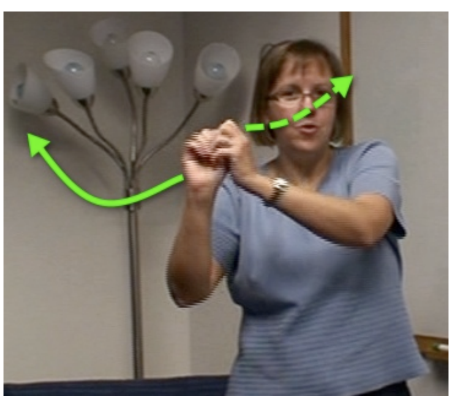


Frame 3*: wide open parking space*

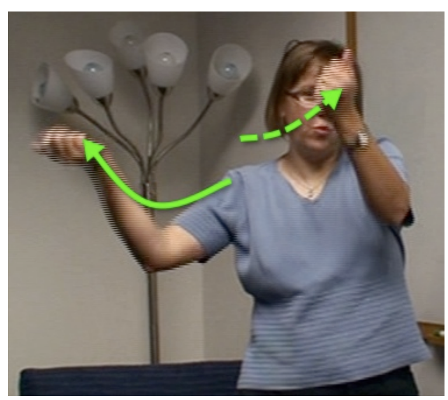


Frame 4. *with people driving in all different directions. There's just a lot more going on, and a lot of other directions that you're talking, think about where you might be making approaches."*

Figure 2. Jane's parking lot metaphor.



Frame 1: Beth illustrating closeness of silly puddy



Frame 2: Beth illustrating a break in the silly puddy

Figure 3. Beth illustrating discontinuity of complex valued functions through silly putty.

Mark provided a different perspective, in terms of how discontinuity is very problematic in the area of physics. He used gestures to explain that, *"... you need to be very careful about your* *models … to make sure that you're staying physical … that you know how to interpret the particular model. You imagine a charge at rest* [put one palm behind the other, with thumbs up, while his hands were facing him as if making a vertical plane] *that you suddenly exert a constant force on* [pushed hands forwards]. *There's a discontinuity there, the force went from 0 to some finite value for a certain amount of time and then drops to zero suddenly*." His words of "*sudden*" and "*drop*" were identical to those used in Jane's hiking metaphor. He elaborated on the fact that sometimes models need to be refined by determining what it would take to make a function smooth. While stating this, he made a motion with cupped hands as if running his hands over a bell-shaped curve.

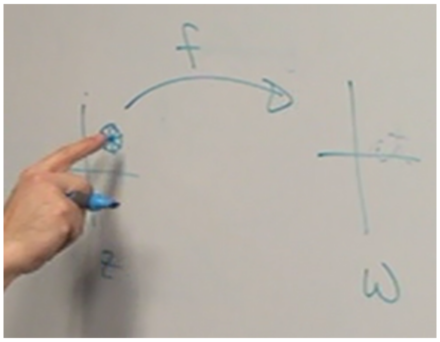
Anton was the only participant to not provide a metaphor for the continuity item. This might be attributed to the fact that he saw no need for a metaphor because as he pointed out, "… *in complex analysis continuity is not the most important one. The most important one is the idea of analyticity. So you don't really think about the continuity.*" Similar comments were made about the question regarding the exponentiation of complex numbers. Everyone except Ricardo explained that he/she had no need to think about raising a complex number to a complex number but expressed a need to consider exponents of the form $\frac{1}{n}$ where n is a whole number, for research purposes. The Cauchy-Riemann equation item did not result in rich geometric responses; everyone but Ricardo jumped into the limit definition and attempted to create a geometric understanding by studying the algebraic representation. While both Ricardo and Jane discussed the role of the symmetries of the Jacobian matrix, Ricardo's explanation found in Soto-Johnson, Oehrtman, and Rozner (2011) appeared to make more geometric connections to the differentiation item, which we discuss next.

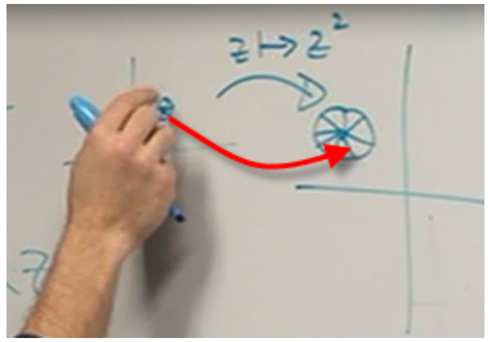
Upon posing the differentiation question, a majority of the participants quickly responded with remarks that one might hear in a multi-variable calculus course related to local behavior and reintroduced the notion of vectors. Such remarks included: "*a derivative is a linearization*", "*it's the local behavior*", "*derivatives only talk about what happens near a point*", "*it's about linear approximations*", "*definitely thinking about it point by point as being a dilation and a rotation*", "*differentiation of any mapping means that a small patch can be approximated by an expansion and a rotation*." These comments were accompanied with drawing a z and w plane with an arrow between them to indicate the mapping. Some participants followed this with points z and z_0 that were close together, while others such as Jane and Luke drew a smaller rectangular or circular grid on top of z_0 . Figure 4 illustrates Jane's drawing of how the rectangular grid on the left gets magnified by $2|z_0|$ and rotated by $\arg(z_0)$; her language was similar to Ricardo's response to the multiplication item. Although Jane drew a rectangular grid, her verbage indicated that in actuality she was visualizing a polar grid. She commented, "*You would draw a grid around the point* (labeled in quadrant I), *I don't know why I am choosing Cartesian, but* (she continues drawing the grid) *so* $f(z_0)$ *is going to go over here* (she plotted it in the second quadrant) *and there might be some curviness to the lines, but I am thinking of the grid as being magnified* (as she drew her image in w-plane). *So like in the real case, I am thinking what happens at a specific point*." The arrow represents her path of rotation at the intersection of two lines at $f(z_0)$ that she followed with her index finger.

Figure 4. Jane's explanation of differentiation.

Luke, like Jane, also started with a rectangular grid but quickly switched over to a polar grid, which he expressed as a bicycle wheel or a pinwheel. Figure 5 illustrates some of the sequencing in Luke's comments. Luke began by mapping the pinwheel through the function $f(z) = z^2$, as he traced the circle in the domain with his right pointer finger (shown in frame 1). In frame 2, Luked used his pincher fingers to represent the "*tiny pinwheel*." Initially Luke thought the only effect of the derivative was a dilation, but after some probing, he commented, that it does actually expand, as he extended his fingers from a closed fist to denote expansion (frame 3) and rotation for a given point. As he said the word rotate he turned his hand as though turning a door-knob (frame 4). He concluded with, "*So the behavior is different at different points, but it's all the same kind of thing. … So for something centered on the real axis, there's not going to be twisting, there's just going to be expansion.*" Ricardo routinely used the same expansion and rotation gestures as Luke, but in the reverse order.

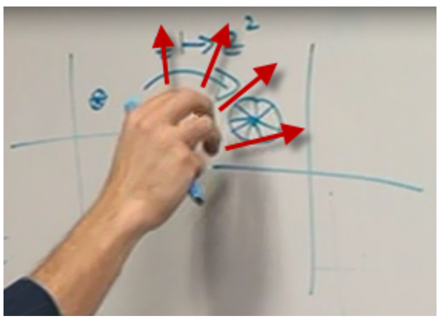
Ricardo started by selecting a point z_0 on the unit circle and taking a "*patch*," which was a small disk around this point. Simultaneously, he remarked, "*it* (referring to the patch) *gets mapped to something that has twice the angle and magnitude is twice as big and then gets spun around*" and he used his right hand to illustrate a turn followed by an extension of fingers to denote the expansion. Ricardo drew a vector to illustrate the radius of the patch and as though his left index finger were the vector, he lifted his finger from the domain to the range and made a circular motion as shown in frame 1 of Figure 7. Since differentiation depends on the point z_0, Ricardo went around the unit circle to denote the different points of reference, as he compared it to a hand on the clock. His left index finger continued to serve as the vector or the hand on the clock. This is illustrated in frame 2 of Figure 6. Ricardo also referenced a spinner, a rigid body, and a turntable in his discussion of differentiation when he segued from his geometric explanation of the Cauchy-Riemann equations into differentiation.

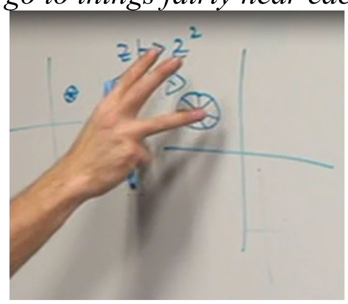




Frame 1: "*It rotates things one way or another depending on the θ. So these little pinwheels...*"

Frame 2: "*All those points near it, at least for a really tiny pinwheel, will look like a factor of two Because all these points are fairly near each other, so they have to go to things fairly near each other.*"

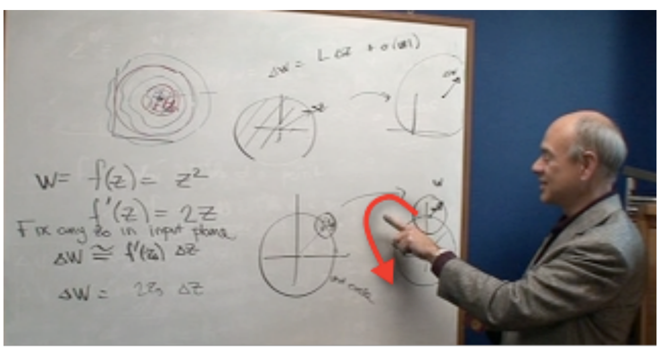


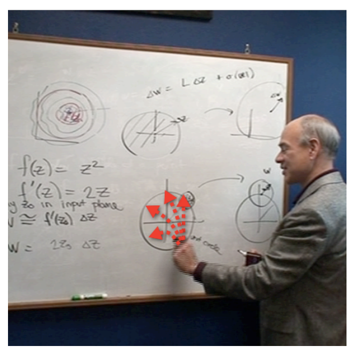


Frame 3: "*So it actually does expand*"

Frame 4: *"and rotate for that point."*

Figure 5. Luke's pinwheel metaphor for differentiation.





Frame 1: "*Small displacement vectors get uniformly approximated by an expansion and a rotation uniformly all at the same time.*"

Frame 2: *"So we get different circles as we go around the unit circle. Dial is going to rotate like a hand on the clock."*

Figure 6: Ricardo's clock metaphor for differentiation.

Discussion

Our research contributes to the minimal existing literature on understanding complex numbers by extending the population beyond high school students and undergraduates and beyond arithmetic and algebraic aspects of complex numbers and equations. While the literature indicates that undergraduates generally possess a static view of the arithmetic of complex numbers (Conner et al, 2007), do not recognize the efficiency of polar representations of complex numbers (Danenhower, 2000), and tend to compartmentalize geometric and algebraic representations (Panaoura et al, 2006), experts appeared to have quite opposite views. This fact is not unexpected since experts have extended and varied experiences; what is of interest to us is how we can use our findings to strengthen students' perceptions of the arithmetic of complex numbers and advanced topics related to complex-valued functions. For example, how can our findings be used to assist students to understand what appears to be "*natural*" for experts become "*natural*" for them. Experts easily recognized the usefulness of working with addition and subtraction with Cartesian form and using polar form for multiplication and division of complex numbers. They also acknowledged how each representation could be connected to the geometry behind the corresponding arithmetic operation. In the arithmetic and differentiation items, the experts easily expressed multiple representations of complex numbers, beyond objects. Complex numbers were fixed vectors, they was a mobile vectors, they were operators or operands, they caused other complex numbers to rotate, dilate, and reflect about the real axis depending on the operation and the location of the complex number. Furthermore, the experts imparted active language to the symbolic representations of complex numbers and in their interpretation of differentiation (i.e., motion, spun, magnification factor, rotation factor, etc.). In summary, the experts appeared to easily and flexibly present and connect symbolic, algebraic, geometric, and verbal explanations of complex numbers and complex-valued functions.

Given we asked our participants to make connections between algebraic and geometric representations of the items, a more significant finding was our experts' use of <u>enacted metaphor</u> (which we did not request as part of the interview protocol) in conveying their perceptions about the content. As such, their reification of the content did not manifest sole through metaphor, but through active descriptions of these metaphors, which illustrated the mathematics. Thus, we propose pedagogical practices allow for activities where students enact their understanding of the mathematics. These enactments may allow learners to invent metaphors that both capture important mathematical concepts accurately and align within their own understandings, experiences, and culture. Another viable way to develop students' use of enacted metaphor is through practitioners' own use of enacted metaphor. The teachers' own dynamic imagery and enacted metaphor may help tap into students' senses beyond hearing and seeing but to include imagining, which may strengthen students' connections between representations. Such practices require that teachers carefully orchestrate and be explicit about their actions and verbiage, since enactments and metaphors can get lost during a lesson due to language barriers, lack of attention, or ambivalence to the role of enacted metaphor in the mathematics classroom. We also suggest that practitioners' offer opportunities for students to reinterpret the practitioners' metaphors into personally meaningful metaphors, that connect to algebraic, symbolic, geometric, and verbal presentations. Technology might also serve as a valuable tool to help students develop enacted metaphor. As Sfard (1994) pointed out, "Because of the tight relationship between the metaphor of an ontological object and the issue of visualization it seems that today's wide accessibility of computer graphs opens promising didactic possibilities" (p. 54).

Further research may include examining how the experts' metaphors represent the mathematics and where they may fall apart, investigating how students use prior knowledge and technology to create and enact their own metaphors, and exploring how creating one's own metaphors may alleviate transfer of complex-variable concepts into other domains such as physics. These veins of inquiry are in our radar for future research.

References

Alcock, L. J., & Simpson, A. P. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 57*(1), 1-32.

Barker, J. M. (2009). *The Tactile Eye: Touch and the Cinematic Experience*. Berkeley: University of California Press.

Conner, M. E., Rasmussen, C., Zanieh, 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.*

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.

Ernest, P. (2010). Reflections on theories of learning. In B. Sriraman & L. English (Eds.), *Theories of Mathematics Education* (pp. 39-49). New York: Springer.

Lakoff , G. & Johnson, M. (1980). *The metaphors we live by*. Chicago: The University of Chicago Press.

Lakoff, G. & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, New York: Basic Books.

Lozano, M. D. (2005). Mathematics learning: Ideas from neuroscience and the enactivist approach to cognition. *For the Learning of Mathematics, 25 (3)*, 24-27.

Miles, M. B. & Huberman, A. M. (1994). *An Expanded Sourcebook: Qualitative Data Analysis*, 2nd Ed. Thousand Oaks, CA: SAGE Publications, Inc.

Nemirovky, R., Rasmussen, C., Sweeney, G., & Wawro, M. (in press). When the classroom floor becomes the complex plane: Addition and multiplication as ways of bodily navigation. *Journal for the Learning Sciences*.

Noë, A. (2004). *Action in Perception*. Cambridge, MA: The MIT 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.

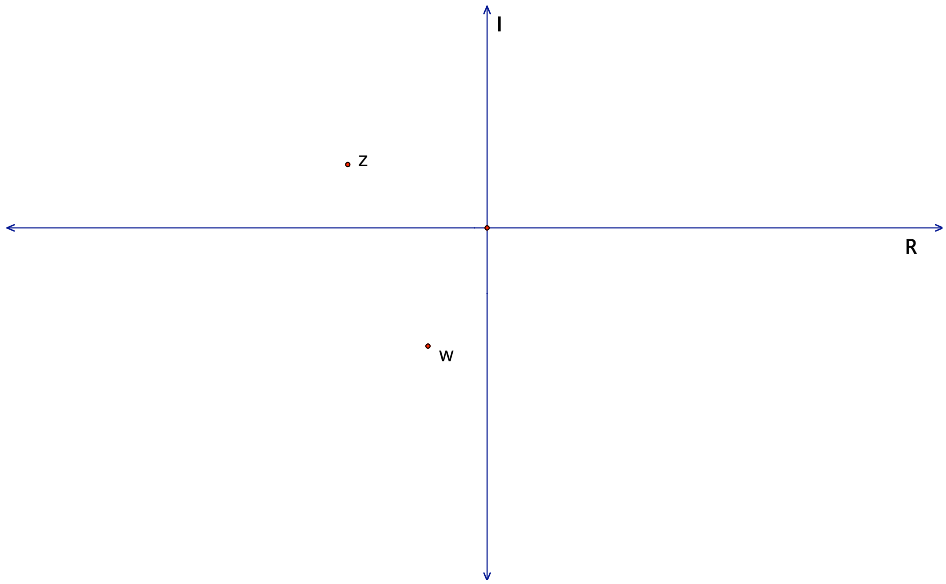
Sfard, A. (1994). Reification as the birth of metaphor. *For the Learning of Mathematics, 14(1)*, 44-55.

Soto-Johnson, H., Oehrtman, M., & Rozner, S. (2011). *Dynamic visualization of complex variables: The case of Ricardo*. Proceeding from the 14th annual conference on research in undergraduate mathematics education, p. 488-503. http://sigmaa.maa.org/rume/Site/Proceedings.html.

Sowder, J. T. (1992). Making sense of numbers in school mathematics. In G. Lenhardt, R. Putnam, & R. A., Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching*. Mahwah, NJ: Erlbaum.

Appendix 1: Interview Questions

1. Below are two complex numbers z and w. Determine and explain how you know where each of the following are located. How do you think of these operations algebraically, geometrically, and in terms of $\mathrm{Re}^{i\theta}$?



a. $z+w$
b. zw
c. $1/z$
d. z^w

2. In calculus, we sometimes use the idea of "tracing the graph of function and not lifting our pencil" to convey the concept of continuity. What geometric representation or explanation might be useful to understand continuity of complex valued-functions?
   a. Consider the function $f(z) = \dfrac{\operatorname{Re} z}{|z|}$. Is there a way to define the function at $z = 0$ in order to make the function continuous? Why or why not?
   b. Consider the function $f(z) = \dfrac{z \operatorname{Re} z}{|z|}$. Is there a way to define the function at $z = 0$ in order to make the function continuous? Why or why not?
3. Give a geometric reasoning as to why it is enough for a real-differentiable function $f(z) = u(x, y) + iv(x, y)$ to satisfy the Cauchy-Riemann equations in order for it to be complex differentiable throughout some ε neighborhood of a point $z_0 = x_0 + iy_0$? Recall that the Cauchy-Riemann equations are: $u_x = v_y$ & $u_y = -v_x$
4. If $f(z) = z^2$, then
   a. What does it mean that $f'(z) = 2z$?
   b. What does it mean that $f'(2) = 4$?
   c. What does it mean that $f'(i) = 2i$?
   d. What does it mean that $f''(z) = 2$?
   e. What does the derivative of a complex function represent? Is it the slope of a line?
5. Sometimes in calculus we can interpret the definite integral of a real-valued function to represent the area under the curve. What geometric representation or explanation might be useful to understand the complex number obtained as an answer to a definite integral of a complex valued-functions?

AN ANALYSIS OF CALCULUS INSTRUCTOR GRADING INCONSISTENCIES THROUGH A SENSIBLE FRAMEWORK

Jana Talley
Jackson State University
Jana.r.talley@jsums.edu

Despite the consensus among mathematics educators that prior knowledge is essential to student success, calculus instructors vary widely in their assessment of prior knowledge errors found on student assignments and exams. This phenomenological study of five calculus instructors at a large research institution investigated the influence that instructor belief systems have on the consistency of grading across instructors. The results showed that the intricacies of instructor sensible systems play a vital role in the assessment of student errors.

Key words: calculus, assessment, prior knowledge, belief systems

Introduction

Anyone who has had an opportunity to work with students taking their first calculus course has probably encountered a variety of student mistakes; none of which have anything to do with the students' understanding of calculus. Whether students fail to manipulate algebraic expressions correctly, forget the values of trigonometric functions at the special angles, or exhibit difficulty sketching simple quadratic functions; assessing prior knowledge mistakes can be quite cumbersome. Within the context of this study the term prior knowledge refers to any skill or understanding a student must possess before entering a first calculus course. Instructors particularly grapple with grading student work when the student demonstrates an understanding of the calculus problem but is unable to successfully complete it due to their deficiencies in prior knowledge. On one hand, the instructor must consider the ability the student has shown in dealing with the topics of calculus. On the other hand, attention must be given to the students' difficulties using the skills taught in previous courses. This contention between the importance of current course objectives and prerequisite skills is settled differently across instructors. Despite the consensus among mathematics educators that prior knowledge is essential to student success (Talley, 2009), variances among calculus instructors' beliefs about prior knowledge in a calculus course yield inconsistent grading of student assignments and exams. Using a phenomenological research design, this study investigated the sensible belief systems of calculus instructors related to the assessment of prior knowledge errors. A sensible belief system is the collection of one's beliefs about a particular construct. A variety of considerations were found to determine instructor grading decisions. The goal of this paper is to report the influences that provoke differences among calculus instructors' grading of student work.

The relevant literature that informs the study is outlined below. Supporting literature includes work concerning the impact of prior knowledge on student learning and student responses to instructor grading techniques. Next, the previous applications of sensible belief systems will be described to lay the foundation for the study's theoretical perspective. The methods and analysis sections clarify how the study was conducted. The analysis section also explains the aspects of the study that propelled the investigation of sensible belief systems; specifically the variances of grading techniques across instructors. Lastly, by applying the sensible systems approach to instructor interview responses, the aspects of student work that give rise to inconsistent instructor grading is revealed.

Relevant Literature

Historically, prior knowledge has been defined in a variety of ways; including some similar to the definition provided above. The terms used to refer to prior knowledge include "prestorage, permanent stored knowledge, prestored knowledge, knowledge store, prior knowledge state, prior knowledge state in the knowledge base, implicit knowledge, or archival memory, not to mention exper[i]ential knowledge, background knowledge, world knowledge, pre-existing knowledge, or personal knowledge" (Dochy & Anderson, 1995, p. 227). Joseph and Dwyer (1984) made use of the term 'entering behavior' and later Cox (2001) employed the term 'probable preparedness' to refer to prior knowledge. A nominal definition (one that specifies the characteristics of a term) is what is used in the current study to solicit conceptions of prior knowledge from each calculus instructor participant.

In many fields of study students have difficulty retaining knowledge from previous course work. Particularly in calculus, the errors that students make have been attributed to prior knowledge and specifically to algebraic misunderstandings in previous research (Edge & Friedberg, 1984; White & Mitchelmore, 1996). One cause of difficulty found by White and Mitchelmore (1996) is students' tendencies to misinterpret the use of variables in calculus problems. They refer to students that manipulate symbols without an understanding of what they are doing as having an 'abstract-apart' concept of variables whereas students who generalize, symbolize, and abstract variables as having an 'abstract-general' concept. They concluded that "a prerequisite to a successful study of calculus is an abstract-general concept of a variable…" (p. 93). Orton's study (1983a) confirms that problems with algebra (in addition to ratio and proportion) hinder calculus students when dealing with differentiation. At Illinois State University, three groups of Calculus I students were studied to determine the factors of success in a first calculus course (Edge & Freidberg, 1984). Edge and Friedberg used regression models to find that for all three groups success in calculus could be predicted by algebraic skills. Research documenting similar situations in calculus, as well as other disciplines, and in varying degrees supports the continued study of this issue. The lack of specific focus on instructors' grading patterns in relation to prior knowledge errors prompted the study outlined here.

The issue of inconsistent grading directly relates to previous research concerning student study habits and intellectual behavior. In The Hidden Curriculum, Snyder (1970) describes the affect that instructional strategies have on the study habits of students. In contrast with what he refers to as the formal curriculum, which traditionally emphasizes deep conceptual understanding of the topics covered in each course, the hidden curriculum is described as the norms that determine successful degree completion which only students understand as insiders of an institution. As an example, Snyder points specifically to a class whose instructor stressed the importance of being creative and engaged in class discussion. However, when presented with the exam, the students found that in actuality they were expected to simply memorize a large portion of their text and regurgitate that information. Students that prevail in environments for which instructor expectations are unclear or vary are known by Miller & Parlett as cue-seekers (Miller and Parlett 1974). Their study of undergraduate science majors revealed that students who carefully gauge the expectations of instructors, despite contradictions to the formal curriculum, perform much better than those who do not read into the hidden curriculum.

The aforementioned research indicates the importance that instructional strategies have on student behavior. Regardless of teacher intentions, the cues sent to our students are indeed received and acted upon. Specifically, the ways in which teachers score exams and assignments are internalized by students and used to tailor future experiences with mathematics learning. Therefore, it is pertinent to the field of mathematics education that we identify those cues. The exploration of factors that influence grading strategies will not only assist students in understanding what is expected of them, but will also allow instructors the opportunity to adjust if those strategies do not align with the intended curriculum.

Furthermore, the increased amount of underprepared students enrolled in college math classes requires more uniform instructional strategies. To ensure fairness, a consensus must be made as to what constitutes important aspects of student work. A first step in that process is to categorize the considerations of instructors. The study reported here begins much needed dialogue by examining the grading strategies of calculus instructors. It not only highlights the differences in opinion that calculus instructors attribute to prior knowledge, but it also outlines other influences that instructors use to determine grades.

Theoretical Perspective

To fully understand the assessment practices of calculus instructors when faced with prior knowledge errors, an examination of each instructor's belief system pertaining to calculus and prior knowledge was required. Belief systems are defined by Phillip (2007) as follows:

> [A belief system is a] metaphor for describing the manner in which one's beliefs are organized in a cluster, generally around a particular idea or object. Belief systems are associated with three aspects: (a) Beliefs within a beliefs system may be primary or derivative; (b) beliefs within a belief system may be central or peripheral; (c) beliefs are never held in isolation and might be thought of as existing in clusters. (p. 259)

To rationalize previously labeled contradictions between teacher beliefs and practices, Leatham (2006) utilized a sensible belief system approach by taking into account a holistic view of teacher belief systems clustered around classroom practice. For example, in a 1997 study of elementary school teachers, Raymond found that the participants' statements concerning beliefs about good teaching strategies did not align with their observed instructional techniques. As a result, these instructors' behaviors were characterized as contradictory to their beliefs. Leatham, and several other researchers (Philip, 2007; Skott, 2009; Speer, 2005; Speer, 2008) opposed this view; citing error in data analysis based on inattention to coexisting beliefs that influenced the teachers' behavior. Because Raymond's study was focused on teacher beliefs about mathematics the interviews did not attend to the plethora of other beliefs that impact the classroom. Beliefs pertaining to time management, instructional resources, and student behavior, just to name a few, are also integral to an instructor's decisions on how to manage a classroom. Leatham went on to re-analyze Raymond's data, concluding that the perceived inconsistencies were in actuality a result of the researcher's narrowly defined use of the term teacher belief. His findings, in fact, showed that the instructors' sensible belief systems exemplified consistencies between teacher beliefs and practices.

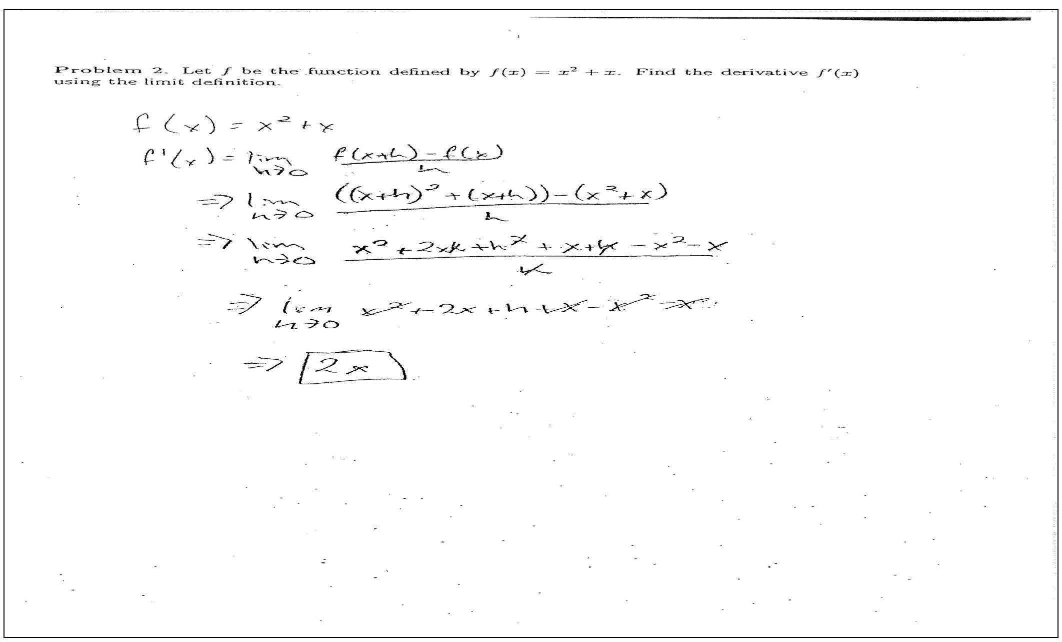
The current study calls upon Leatham's (2006) assertion that, as observers, the perceived beliefs of another are assumed consistent or contradictory based on our own perspectives. "The sensible system framework attempts to minimize these assumptions" (p. 95). He considered the entire system when analyzing teacher behavior that appeared to contradict teacher beliefs. Rather than conclude that the teacher was conflicted when observed instructional behavior did not align with a stated belief, Leatham resolved that there were other beliefs existing within the teacher's belief system that took precedence at the time of the perceived contradictory action. An adaption of the sensible system framework is used here to rationalize the variances among assessments of calculus student errors. Further details of this adaption are reserved for the analysis section.

Methods

A qualitative research design was used to investigate instructor perspectives on prior knowledge. A phenomenological approach was taken to uncover instructor views of how prior knowledge skills influence student performance in calculus and how prior knowledge errors influence instructor judgment of student understandings. Therefore, interviews were designed to first identify each instructor's definition of prior knowledge and using that definition their grading techniques were explored. Five calculus instructors at a large Midwestern research institution were individually interviewed. Each participant was a faculty member who had taught a 120-student lecture style calculus course within the last five years of being interviewed. These courses were structured in such a way that the students met for lecture with the faculty instructor for a one-hour lecture three times per week. The students also met with a graduate teaching assistant one hour each week for a recitation style course. Under the direction of the faculty instructor the graduate assistant was normally required to demonstrate additional examples during recitation, collect and grade homework, hold office hours, and assist in the grading of exams.

The interviews were two-pronged consisting of a traditional interview component and a task-based component. The items included in the first component of the interview centered around three main issues. The participants were first asked to provide a definition of prior knowledge in a calculus course by listing the skills students need to be successful in a first calculus course. They were also asked to describe how important prior knowledge is in a calculus course. Lastly, the instructors explained how they approach prior knowledge errors when grading.

The second component of the interviews was task-based, requiring the participants to score a selection of student exam questions. The exam questions were collected during a pilot study from students at the same institution during the previous semester. The use of student exams from a previous semester ensured that the instructors were unfamiliar with the students to prevent bias in the scoring process. These student error examples (SEEs) were chosen to reflect questions commonly seen on the university's exams.


Problem 2. Let f be the function defined by $f(x) = x^2 + x$. Find the derivative $f'(x)$ using the limit definition.

$f(x) = x^2 + x$

$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$

$\Rightarrow \lim_{h \to 0} \frac{((x+h)^2 + (x+h)) - (x^2 + x)}{h}$

$\Rightarrow \lim_{h \to 0} \frac{x^2 + 2xh + h^2 + x + h - x^2 - x}{h}$

$\Rightarrow \lim_{h \to 0} x^2 + 2x + h + x - x^2 - x$

$\Rightarrow \boxed{2x}$

Figure 1. Student Error Example 1 (SEE 1). One of nineteen SEEs presented to interviewed instructors.

Specifically, the question types fall into the following seven categories:

- Find the derivative using the limit definition
- Find intervals of continuity
- Find the derivative using rules

- Find equation of a tangent line
- Use techniques of implicit differentiation
- Solve the related rates
- Solve maximum/minimum applications

Nineteen SEEs were presented to the interviewed instructors. Out of a given point value the instructors were asked to score each SEE. They were also asked to classify the error(s) the student made, if any, as calculus errors or prior knowledge errors. In the case that an error was classified as a prior knowledge error the instructor identified the type of prior knowledge error whenever possible. Lastly, the instructors commented on how their scoring decision was made.

Analysis

The first component of the interviews prompted the instructors to give their general opinion about prior knowledge and its influence on the Calculus I course. These responses were used to identify themes among the instructors that would help to characterize a set of abilities, understandings, and/or skills that faculty expect their students to possess prior to entering a first calculus course. Preliminary analysis of interview transcripts and expanded field notes uncovered four areas that instructors focused on when describing their expectations of students entering their Calculus course: algebra, trigonometry, an understanding of functions, and scholarly enthusiasm. Of these four types only the first two were mentioned by all five of the interviewed instructors. Proficiency in algebra was spoken of as the foundation needed to develop student understanding of the concepts of calculus. Additionally, a student with little or no understanding of trigonometric functions was seen by these calculus instructors to be at a disadvantage in the course.

Of central importance to this report, analysis of the task-based component of the interviews involved a comparison of the participants' assessments of the nineteen SEEs. Particular attention was given to the instructors' scoring of student work. The difference between the lowest and highest assigned score was identified for each SEE. Interestingly, several of the error examples revealed scores that ranged from below 70% to above 80%. This was alarming because at the participants' institution a score below 70% was considered a failing grade for majors and a score greater than 80% was considered above average. Therefore, these seven SEEs were later labeled as having wide score ranges. The following table outlines the questions that were identified as having wide score ranges. It should be noted that the wide score ranges were found only among four questions types: (a) Find the derivative using the limit definition, (b) Find intervals of continuity, (c) Find the derivative using rules, and (d) Find equation of a tangent lines.

Table 1.

Instructor Scores of Wide Score Ranges

Find Derivative Using Limits		Find Intervals of Continuity		Find Derivative Using Rules	Find Equation of Tangent Line	
SEE 1 (20pts)	SEE 2 (10pts)	SEE 3 (10pts)	SEE 4 (10pts)	SEE 5 (10pts)	SEE 6 (15pts)	SEE 7 (15pts)
Prof. E 10	Prof. C 5 – 6	Prof. E 3	Prof. F 2	Prof. F 5 – 6	Prof. E 10	Prof. E 10
Prof. F 10 – 12	Prof. D 6	Prof. F 3	Prof. E 3 – 4	Prof. G 6 – 7	Prof. G 10	Prof. F 12
Prof. G 12 – 15	Prof. G 7	Prof. D 5	Prof. G 5	Prof. E 7	Prof. F 10 – 11	Prof. G 12
Prof. D 15	Prof. F 8 - 9	Prof. C 6	Prof. C 8	Prof. D 7	Prof. D 12	Prof. C 12 – 13
Prof. C 18	Prof. E 9	Prof. G 8	Prof. D 8	Prof. C 8	Prof. C 14	Prof. D 13
Score Range 50 – 90%	Score Range 50 – 90%	Score Range 30 – 80%	Score Range 20 – 80%	Score Range 50 – 80%	Score Range 66 – 93%	Score Range 66 – 86%

Note: If an instructor responded with more than one possibility for a score it is denoted within the table as 'lowest possibility - highest possibility'.

To investigate this phenomenon of wide score ranges I looked to the work of Leatham (2006) in the study of teacher belief systems. He found that perceived inconsistencies between instructor beliefs and practices could be explained by taking a more holistic look at the teachers' beliefs about pedagogy. In the same vein, I viewed the inconsistent responses from the group of instructors to be a function of the belief systems in play by each individual professor. In this study the unit of analysis is the set of score assignments of each SEE. For each error example with a wide score range, the instructor comments were compiled and analyzed to determine how their individual sensible systems impacted grading decisions and more specifically if and how their position on calculus, prior knowledge, and assessment supported their scorings.

Findings

Seven of the nineteen SEEs were identified as having wide score ranges. To demonstrate how the sensible system framework was applied, the two most interesting cases will be discussed here. SEE 1 and SEE 2 both required the student to find the derivative of a given function using the limit definition of the derivative. As displayed in Table 1, the instructor scoring in regards to highest and lowest scores are reversed in these cases. In SEE 1, Professor Crumbliss gave the highest score and Professor Edwards the lowest score. However, Professor Edwards gave the lowest and Professor Crumbliss the highest for SEE 2. Also, the scores of the other three instructors maintained this inverted pattern of scoring. Because of the unique relationship between the score ranges of SEE 1 and SEE 2, the influence of specific aspects of student work on scoring decisions is most clearly exemplified through their examination. Therefore, a detailed account of how the sensible system framework was applied to these two cases will be provided here by comparing the grading rationales of Professor Crumbliss and Professor Edwards. Similar findings were uncovered in each of the remaining six cases of wide score ranges.

In SEE 1, (shown in Figure 1 above) Professor Crumbliss gave the highest score and Professor Edwards the lowest score. Through a holistic lens of each instructor's perspective on teaching, assessment, and prior knowledge, in particular, the differences in scoring were understood. SEE 1 was scored out of twenty points. Professor Crumbliss assigned the highest score of 18/20. In his opinion, the algebraic mistake made by the student was of little importance compared to the student's ability to demonstrate conceptual understanding. In his interview he commented that:

"The most frustrating one is when someone comes up with the equation of a tangent line and it's just perfect and then they go and they simplify it more and there is an algebra mistake there and it's just, I will usually give that full credit because I think that stuff is irrelevant. You know it's more important that they get the concepts."

"There are examples where the calculus is all true and people organize things and make mistakes. I'm not worried, I'll usually circle it and write some type of comment and give them nine out of ten."

<u>Sensible System</u>: Professor Crumbliss felt he could assign an above average score for two main reasons. The first is that he can clearly see that the student understands the process and was able to set everything up correctly. Also, when everything else is correct, the algebra mistakes were seen as irrelevant. Therefore, such errors warranted only minimal point deduction, if any at all.

The lowest score given to SEE 1, 10/20, was assigned by Professor Edwards. He described the students' inability to complete each step of the problem correctly as problematic. Specifically, he made the following assertions:

"So if a student shows that they have some conceptual understanding I do give them some credit. But calculus is about calculations and you need to get the calculations right. So even if it's a question of a deficiency in prior knowledge it's still something [the student is] responsible for."

"…So they wrote down the derivative correctly and wrote down the correct limit and put in the function and expanded the functions. So the whole problem was in the manipulating…They did that incorrectly and then they got the right limit. So I'd give maybe 10 points [out of 20]."

<u>Sensible System</u>: Professor Edwards explained that "calculus is about doing calculations". He views algebraic manipulations as part of the work of calculus. Partial credit (half in this case) is assigned when students demonstrate good work, which he was able to identify in SEE 1.

SEE 2, scored out of 10 points, also required that the student find the derivative using the limit definition. In contrast to SEE 1, Professor Edwards assigned the highest score of 9/10. He attributed the student's mistake to a small lapse in focus. His response to the error was:

"This student has a very clear idea of what's going on. So there are some very minor slips. So they would get almost full points…maybe 9 points. Interviewer*: Are those minor slips in calculus or prior knowledge.* Participant*: Oh umm, concentration. I mean the student is skilled enough I think they made a quick error."*

"Yeah that's something I wouldn't be so harsh with. You just assume they were rushing. They did everything correctly up to then it seems. So yeah, that's just a slip of concentration."

Sensible System: Professor Edwards considered slips or lapses in concentration to be somewhat insignificant. He was able to recognize that the student "has a clear idea of what's going on" and therefore deserves close to full credit.

Professor Crumbliss, on the other hand, gave the lowest score range of 5/10 to 6/10 for SEE 2. He considered the student's errors to be quite serious, specifically in terms of conceptual understanding. He stated:

"Its not clear that they are comfortable using the algebra or that they have any clue as to what's going on conceptually."

"This is a hard one because they're making really horrible errors like two squared is two… I don't know maybe 6 and a conversation or maybe five and a conversation."

Sensible System: A major concern for Professor Crumbliss here is that the student could not demonstrate their ability to use algebra or calculus. Conceptual understanding, which he views as most important, is not exhibited in this case.

This snapshot of the sensible system framework provides an avenue for understanding variances in instructor grading patterns. This examination of the instructors' perspectives on the types of errors and conceptual skills demonstrated in student work proved to be especially insightful. The sensible system framework applied to SEE 1 revealed that each instructor was thoughtful in their considerations despite the varied level of scores assigned. Professor Crumbliss was concerned with the ability of the student to demonstrate conceptual knowledge regardless of the existence of algebra mistakes. Conversely, Professor Edwards' attention was given to the students work as a whole and his grading decisions hinged on how well the student worked through prior knowledge skills in addition to calculus procedures. As shown here, the existence of a prior knowledge error has a different meaning for each instructor. For Professor Crumbliss, prior knowledge errors are stumbling blocks that can be addressed by small point deductions. On the other hand, Professor Edwards sees those same mistakes as indicators that the student cannot be successful in a calculus course.

It is also interesting that Professor Edwards and Professor Crumbliss classified the errors in SEE 2 differently. Professor Edwards attributed the mistakes to sloppiness and Professor Crumbliss viewed them as evidence of conceptual misunderstanding. These differences in evaluation of the student's work clearly called upon different aspects of the instructors' sensible belief systems. Though Professor Edwards places heavy emphasis on the importance of student demonstration of both understanding of calculus and understanding of prior knowledge, when he perceives a student error to be a consequence of sloppiness or a minor slip in concentration, he is willing to overlook missteps and assign an above average score. For SEE 2, the assertion that conceptual understanding most heavily influences Professor Crumbliss' grading decisions is confirmed as he assigned a score of 5 out of 10 after determining that the student had no clue as to what was going on conceptually.

Implications

The importance of prerequisite skills in mathematics courses has been well documented. Particularly at the college level, students need a foundation of prior knowledge to navigate through mathematics requirements and specialized courses in their respective fields of study.

However, the increase in college and university enrollment as of late has flooded post secondary classrooms with underprepared students who lack sufficient prerequisite skills. This influx of underprepared students does not exempt university professors from attending to course objectives designed to build upon the much needed prerequisite courses such as algebra, trigonometry, and geometry. The insights this study provides with respect to instructor belief systems should be carefully considered as mathematics educators develop methods for providing instruction to students lacking necessary prior knowledge skills.

Recent increases in class sizes have also fueled trends towards uniformity across multiple section courses like calculus. Mathematics departments are now looking to provide students with consistency in various aspects of the classroom experience; especially in grading policies. These efforts towards fairness should be tempered with an understanding of the decisions instructors make when assessing their students. The results of this study provide a backdrop for administrators and faculty who manage the coordination of multiple-section classes as they consider the grading practices to be incorporated into redesigned curriculums.

Further Research

As this study attended to instructor views concerning prior knowledge and calculus grading, many questions about the instructor were not explored that may shed light on why they hold particular beliefs about mathematics educational practices. Perhaps it would be useful to develop instructor profiles that include characteristics such as research areas (pure or applied), secondary and post secondary schooling experiences (private, public, or U. S.), or institution type (research or teaching). These profiles could then be used to determine if and what correlations exist between grading techniques and instructor characteristics.

As noted in the analysis section, only four of the seven question types yielded wide range scores. What aspects of these types of questions might have prompted such disagreement across instructors? One conclusion could pertain to the fact that the SEEs from the remaining three question types (Implicit Differentiation, Related Rates, and Maximum/minimum Applications) received below average scores across the board. These three types of questions are usually presented during the latter half of the semester and students have less time to become proficient in those processes. Therefore, all of the students tend to perform poorly which in turn leaves fewer opportunities for instructors to vary in their assessment of student work. Despite the likelihood of this assumption, a more structured investigation is necessary to understand the variances of the types of questions presented here and for those in other mathematical subjects as well.

References

Cox, W. (2001). On the expectations of the mathematical knowledge of first-year undergraduates. International Journal of Mathematical Education in Science and Technology, 32(6), 847-861.

Dochy, F. J. R. C., & Alexander, P. A. (1995). Mapping prior knowledge: A framework for discussion among researchers. European Journal of Psychology of Education, 3, 225-242.

Edge, O. P., & Friedberg, S. H. (1984). Factors affecting achievement in the first course in calculus. The Journal of Experimental Education, 52(3), 136-140.

Joseph, J. H., & Dwyer, F. M. (1984). The effects of prior knowledge, presentation mode, and visual realism on student achievement. The Journal of Experimental Education, 52(2), 110-121.

Leatham, K. R. (2006). Viewing mathematics teachers' beliefs as sensible systems. Journal of Mathematics Teacher Education, 9, 91-102.

Miller, C. M., & Parlett, M. (1974). Up to the mark: A study of the examination game. London: Society for Research into Higher Education.

Philipp, Randolph A. (2007). Mathematics teachers’ beliefs and affect. In F. Lester (Ed.), Second Handbook of Research on Mathematics Teaching and Learning. Reston, VA. National Council of the Teaching of Mathematics.

Philipp, Randolph A. (2007). Mathematics teachers’ beliefs and affect. In F. Lester (Ed.), Second Handbook of Research on Mathematics Teaching and Learning. Reston, VA. National Council of the Teaching of Mathematics.

Skott, J. (2009). Contextualising the notion of 'belief enactment'. Journal of Math Teacher Education, 12, 27-46.

Snyder, B. R. (1970). The Hidden Curriculum. New York: Alfred A. Knoff.

Speer, N. M. (2005). Issues of methods and theory in the study of mathematics teachers' professed and attributed beliefs. Educational Studies in Mathematics, 58, 361-391.

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, 218-267.

Talley (2009). Calculus instructors’ responses to prior knowledge errors. Unpublished Dissertation, The University of Oklahoma, Norman, OK.

FOR EDUCATIONAL COLOR WORK: DIAGRAMS IN GEOMETRY PROOFS

Allison F. Toney
University of North Carolina Wilmington

Kelli M. Slaten
University of North Carolina Wilmington

Elisabeth F. Peters
University of North Carolina Wilmington

Historically grounded in Oliver Byrne's reworking of Euclid's Elements, and based on a student-generated proof, we investigate the use of coloring to enhance geometry proofs. Charlotte Knight, an undergraduate mathematics major enrolled in a modern geometry course, regularly employed coloring techniques as a tool in her proof-writing. We met for a single semi-structured, task-based interview to discuss Charlotte's use of coloring in her organization and understanding of geometry proofs. Results indicate that Charlotte's use of diagrams is closely related to her construction of a proof. In particular, her use of color serves several purposes: (1) as an organizational tool to connect her diagrams to the content of her proofs, (2) to enhance her understanding of the proof she is writing, and (3) to illustrate relationships within her diagrams and proofs. We believe this small study has particularly interesting pedagogical implications at the post-secondary level as well as for K-12 mathematics instruction.

Keywords: modern geometry, proofs, diagrams, color

Background

In 1847, Oliver Byrne published his reworking of Euclid's *Elements*, in which he used colored diagrams so extensively that the visual representations were inseparable from the proofs they were intended to support. Published during a period when geometers had their attention focused on non-Euclidean investigations, Byrne's work was not taken seriously, and was "regarded as a curiosity" (Cajori, 1928, p. 429). However, Byrne did not intend his work for mere entertainment. Instead, he proposed that the book enhanced pedagogy by appealing to the visual and encouraging retention of the ideas. He suggested that by communicating Euclid's ideas through a colored, visual means, instruction time could be used more efficiently and student retention is more permanent (Byrne, 1847).

Students' transition to formal proof is a well-documented area of research in mathematics education (e.g., Moore, 1994; Selden & Selden, 2003; Weber 2001). However, students' use of diagrammatic representations to support their arguments is still an emerging field of research at the post-secondary level. Where there is considerable research available about calculus students' use of visual representations (e.g., Hallet, 1991; Tall, 1991; Zimmerman, 1991), there is little research available about students in advanced undergraduate mathematics.

Additionally, the National Council of Teachers of Mathematics (NCTM, 2000) asserts that creating and using representations is an essential component to mathematical understanding. As a result, the use of visual representations in K-12 mathematics (in particular, K-12 geometry) is well-documented (e.g., Christou, Mousoulides, Pittalis, Pitta-Pantazi, 2004; Hanna, 2000; Ye, Chou, &Gao, 2010). The Conference Board of the Mathematical Sciences (2000) recognized a collegiate geometry course as a typical component of teacher preparation curricula, where students gain essential skills in visualization, “understanding the nature of axiomatic reasoning”, and “facility with proof” (p. 41). However, little research exists concerning students’ proving in undergraduate modern geometry courses.

In his research investigating students’ use of visual representations in an introductory analysis course, Gibson (1998) found that students implement diagrams to (1) understand information, (2) determine the truthfulness of a statement, (3) discover new ideas, and (4) verbalize ideas. Using diagrams to understand information relates to students’ use of diagrams to aid in their comprehension. Determining truthfulness pertains to student-generated diagrams as a means to confirm or refute mathematical statements. Students discover new ideas when they use diagrams to “get unstuck” and gain insight on the direction of their proof. Using diagrams to verbalize ideas relates to students drawing diagrams to reduce or minimize their cognitive load.

Yestness and Soto (2008) used Gibson’s results to frame their study of 7 students who used diagrams in the development of their understanding of abstract algebra concepts. They found that students most commonly employ (1) and (4) in their diagramming. In particular, they discussed students who explained that their drawings were merely for personal use and not for proof or explanation. However, when asked to explain their proof, many drew a diagram to support their explanation.

The primary goal of this small research study was to bring Byrne’s work into an investigation of how students in an undergraduate modern geometry class use diagrams as proof-writing tools. In particular, we noticed a growing number of students employing the use of color to support their diagrams in our advanced undergraduate mathematics classes. We used a framework proposed by Gibson (1998) and reinforced by Yestness and Soto (2008) to guide our small phenomenological research study into a single geometry student’s use of color-enhanced diagrams as a proof-writing tool. The question guiding our research is: *What is the nature of students’ use of color as a proof-writing tool in college geometry?*

Methods

The setting for our research took place at a medium-sized public university in the southeastern United States. We purposefully identified Charlotte Knight, an undergraduate mathematics major with a concentration in teacher licensure, as a participant because of a “colored” proof she provided on an in-class exam in a modern geometry course (Patton, 2002). Upon further investigation we found that Charlotte regularly employed coloring techniques in her proof-writing. Very similar to the proofs Oliver Byrne presented in his reworking of Euclid’s *Elements*, we were curious about Charlotte’s reasoning.

We met with Charlottefor a single 75-minute semi-structured, task-based interview. The audio-recorded interview focused on two main components. We first presented her with one of the original colored proofs she submitted, in which she correctly proved that the diagonals of a parallelogram bisect each other. We asked her to recount the process she followed while writing the proof. In the second part of the interview, we gave her the Pointwise Characterization of Angle Bisectors Theorem:

> Let□, □, and□be three noncollinear points and let□be a point in the interior of∠□□□.
> Then P lies on the angle bisector of∠□□□if and only if□□,□□=□□,□□.

She had previously completed this proof for homework in her modern geometry class, and we asked her to work through a proof again with us in the interview. We had colored pens available on the table for her to use.

We used constant-comparative methods of analysis as outlined by Corbin and Strauss (2008). That is, using the transcription of the interview, we systematically open and axial coded the data to identify emergent themes in Charlotte's interview, while regularly revisiting the theory identified in Gibson (1998) and supported by Yestness and Soto (2008).

To ensure accuracy, as we were drawing conclusions from the interview, we followed up again with Charlotte (Creswell, 2003). We shared our findings with her and asked for clarification and suggested revisions.

Results

Charlotte spent significant time in the interview describing how and why she used color to enhance her proofs. She also used color extensively in the proof we asked her to construct.Results indicate that all four aspects of diagramming offered by Gibson (1998) and supported by Yestness& Soto (2008) are apparent in Charlotte's colored proofs.

We found significant support for each of Gibson's categories. In particular, determining truthfulness of statements and writing out ideas stood out as being prevalent in Charlotte's proving process.She used color to confirm or refute ideas and pathways she took. She also used it to reduce her cognitive load - it was less mentally taxing to use color over symbols and words. Including color served to help her sort and organize relationships, which she then used to write out her proofs.

Additionally, we found that Charlotte used colors in two primary ways: (1) as a managerial tool to understand the theorem, and (2) as an organizational tool to connect her diagrams to the content of her proofs.

Color as a Tool to Understand the Theorem

Charlotte regularly used color to indicate direction in a theorem. She primarily did this when proving "if and only if" theorems, saying she was uncomfortable with these because she had difficulty keeping track of which direction she was proving, what information she could assume, and what she was trying to show.

In her proof of the Pointwise Characterization of Angle Bisectors, Charlotte used a 2-color scheme. All information in the necessary direction was designated green and all information in the sufficient direction was designated blue. She then constructed and colored a diagram to reflect this information. As a result, to Charlotte, the statement of the theorem changed from "□ lies on the angle bisector of $\angle$□□□ if and only if □□, □□ = □□, □□" to "green if and only if blue" (see Figure 1).This served to help her reduce the cognitive requirements put forth with the 2-direction theorem, as well as to verify her belief in the truthfulness of the statement of the theorem. It also aided her understanding of the information required to construct a proof. She said, "I'm not as familiar with this picture…so I needed to keep referencing back and forth here and so I needed to know…it's kind of like a help to know where I'm going and it's, it's a reference." She said using the color helped her stay organized, understand the theorem, and stay on track with her proving goals:

> This helps me remember which direction I'm going, 'cause all the green stuff is what I knew from the first half of the statement… I put all of that in green and then if and only if…I put in blue on this one so that I knew my directions.

As she continutedto construct the proof in the interview, Charlotte added a second layer of coloring – one in which she used color to understand and manage the mathematical content.

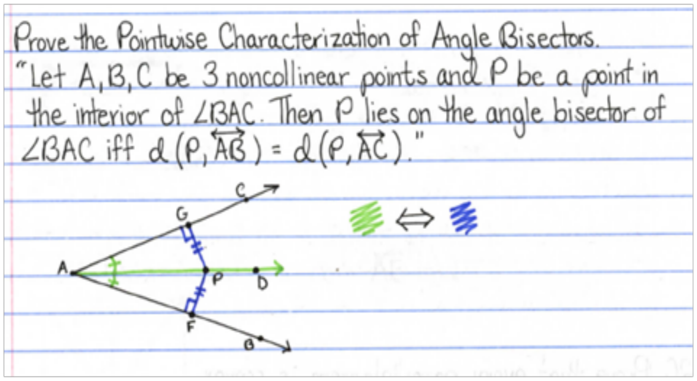


Figure 1. Charlotte's reworking of the statement of the Pointwise Characterization of Angle Bisectors theorem to "green if and only if blue", as put forth in a class homework assignment.

Color as a Tool to Manage Proof Content

Charlotte initially reworked the Pointwise Characterization of Angle Bisectors into "green if and only if blue." Once she began to prove the theorem, however, she went on to include additional colors to guide her through the mathematical content of the proof. At this level, Charlotte continued to use colors to write out her ideas about the proof, and to understand the concepts within her proof. But she also used this coloring technique to gain insight about her proof:

> Then I can see that I have this one and I have this one and I have this one, um, and I don't have anything over here. Like I don't have AG, I don't know anything about AG so there's no colors or label, there's no nothing. I don't know anything about FA. What I do know is all in color, so it kind of helps me know well this is what I have to work with, because I don't want to go try to prove FA and FG, I don't have anything to work with to get there, so it helps that I have the purple angles here to say these are right…I don't think I used anything that wasn't related to color in some way. Like I'd never talked about just the segment FA, you know what I'm saying? I talked about segment AP, but I gave it a blue squiggle.

In the proof she was asked to recount, where she had proved the diagonals of a parallelogram bisect each other, Charlotte employed a 3-color scheme. She used these colors in a way in which the diagram was inseparable from the proof it was intended to accompany. She colored the angles in a way that corresponds to the underlined colors in her proof (see Figure 2). In describing this proof, she used the language "purple is congruent to purple," "orange is congruent to orange," "pink is congruent to pink," and "green is congruent to green." That is, the colors essentially replaced the alphabetical identifiers and this is how Charlotte navigated her proof:

> I needed to look at, like, labeled the purple angles and then I underlined them for both so I knew purple was done…now which one is similar to the purple ones…to the orange ones and then I have pink and green left, well pink and then which one is similar to pink? Green. So that's how that, that's how that went.

In recounting this proof, she spoke primarily of using colors to organize her ideas and understand the information required to write the proof.

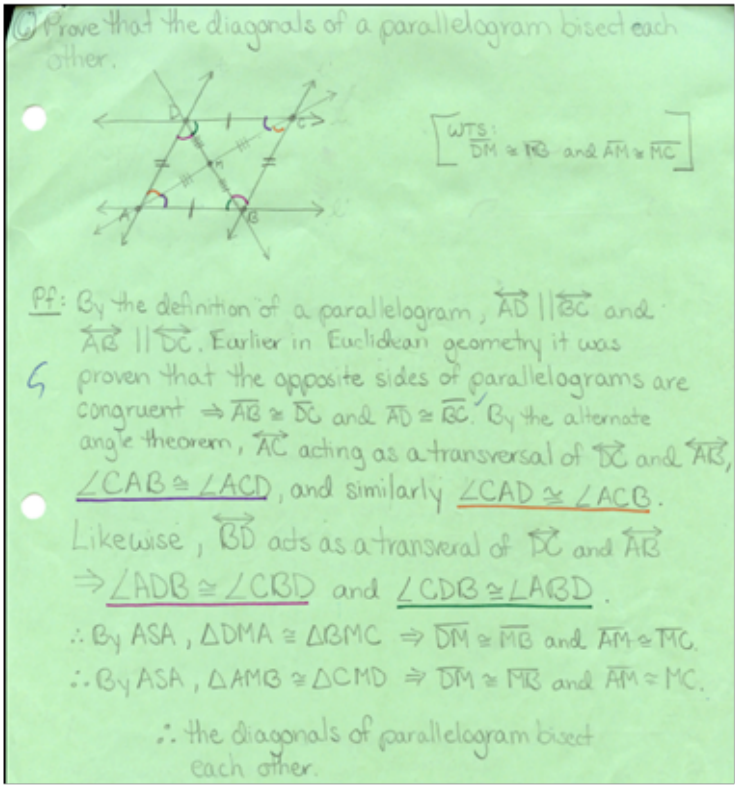


Figure 2.Charlotte's colored proof of the statement that the diagonals of a parallelogram bisect each other.

Conclusions

We found Charlotte strong evidence that Charlotte moved in, around, and through the 4 categories, while moving through three distinguishable stages: (1) pre-proof, (2) proof, and (3) post-proof.

In the pre-proof stage, Charlotte mapped out her plan of attack.In follow-up member-checking she commented on this using a puzzle metaphor. She said she was getting a feel for the big picture, sorting the pieces, and trying to figure out where different groups of pieces will go. She said "when you have the ocean, it's just a bunch of blue pieces. You know what the ocean looks like but you want to scream because the pieces won't just fit where ever." She started working on her proof by first determining truthfulness – drawing and coloring a picture to confirm that she actually believed the theorem to be true. She exchanged the words and symbols for colors to free up cognitive space to focus on her understanding of the theorem and the concepts she "has in her toolbox." At the pre-proof stage, she used color to collect and group information in order to understand the concepts put forth in the statement of the theorem and to gain insight about the direction she needs to take.

Once she began constructing her proof, what she would ultimately deem appropriate to "turn in for a grade," she referred back to her puzzle metaphor. She characterized this as the point where she actually "sat down and figured out how the ocean pieces fit together to complete the puzzle." The primary difference in the proof stage from the pre-proof stage is the nature of determining truth and the disappearance of understanding ideas. We began to see Charlotte reflecting on the truthfulness of her own claims and statements, and using colors to aid in these reflections. We also donot see Charlotte working to understand information. It may be the case that we don't see this here because its presence might indicate that Charlotte was not ready to move out of pre-proving.

Once she had written a proof she believed she was satisfied with, Charlotte stepped back and reflected on her process. In the post-proof stage, Charlotte reflected on the "correctness" of her proof. Here, writing out ideas served as more of a check list and determining truthfulness was about self-reflection of the truthfulness of the claims or statements she made in her proof.

Future Work

We see several threads for future work.First, we want to look deeper into how Charlotte uses color at each stage of her proofs. We would also like to expand our participant pool and investigate these conjectures with other students.

Not addressed in this research report, but an important part of Charlotte's interview and follow-up, is a question about what constitutes a rigorous mathematical proof. Charlotte was verbal about her discomfort with the idea of a color-only proof and she spent significant time trying to reconcile what constitutes a "correct proof" .For her own use, she felt it was sufficient to have a colored proof. That is, she indicated a statement such as "If blue is congruent to blue, and purple is congruent to purple, then red is congruent to red" might be sufficient for her notes. However, she asserted that without shared meaning, a proof such as this would not be sufficient – this would not be a correct proof for "mixed company:"

> So I felt like I needed to go back and write, I need to write what I'm talking about. 'Cause I don't feel like, I don't know, you would never look at a proof in a textbook that has purple squiggles…So I needed to be professional about it and I mean it's a test and I needed to do it right…Then, I'm turning in homework to a professor, so I just wrote it out how I expected she would want to see it…I don't mean for that to sound bad but…I wouldn't expect my students to write their homework all in different, completely different ways using completely different methods and we have to like sit and interpret and be the detective.

We think this connects in an interesting way to Byrne's reworking of *The Elements* not being taken seriously and could lead to further investigation about the nature of students' understanding of mathematical argument.

Pedagogically, we find Charlotte's case to be particularly interesting. Byrne asserted that using color-coded proofs allows students to see the key parts of the argument rather than having to mentally connect what the letters refer too, and thus reducing opportunities for confusion. Charlotte's interview provides substantial support for this argument. Extending this research to include other participants who utilize diagrams (and, in particular, utilize colored diagrams) may shed light onto how to reform instruction accordingly.

References

Byrne, O. (1847). *The first six books of the elements of Euclid in which colored diagrams and symbols are used instead of letters for the greater ease of learning.* London: William Pickering.

Cajori, F. (1928).*A history of mathematical notations, volume 1: Notations in elementary mathematics*. London: Open Court Company.

Christou, C., Mousoulides, N., Pittalis, M., & Pitta-Pantazi, D. (2004).Proofs through exploration in dynamic geometry environments. In M. J.Hoines& A. B. Fuglestad (Eds.) *Proceedings* *of the 28th Annual Meeting of the International Group for the Psychology of Mathematics Education*, Bergen, Norway, July 11 - 14, 2004.

Conference Board of the Mathematical Sciences (2000).Recommendations for high school teacher preparation.*The Mathematical Education of Teachers.*(pp. 37-46). Providence, RI: American Mathematical Society.

Corbin, J. & Strauss, A. C. (2008).*Basics of Qualitative Research*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks: Sage.

Gibson, D. (1998). Students' use of diagrams to develop proofs in an introductory analysis course. In A.H. Shoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education. III* (pp. 284-307). Providence, RI: American Mathematical Society.

Hallet, D. H. (1991).Visualization and calculus reform. In W. Zimmerman & S. Cunningham (Eds.) *Visualization in teaching and learning mathematics: MAA notes number 19* (pp. 121-126). Washington, DC: Mathematical Association of America.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics, 44*(1), 5-23.

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: NCTM.

Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods.* Thousand Oaks: Sage.

Selden, J. & Selden, A. (2003). Validation of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, *34*(1), 4-36.

Tall, D. (1991). Intuition and rigor: The role of visualization in the calculus. In W. Zimmerman & S. Cunningham (Eds.) *Visualization in teaching and learning mathematics: MAA notes number 19* (pp. 105-119). Washington, DC: Mathematical Association of America.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101-119.

Ye, Z., Chou, S. C., Gao, X. S. (2010). Visually dynamic presentation of proofs in plane geometry.*Journal of Automated Reasoning*, *45*, 213-241.

Yestness, N. & Soto, H. (2008).Student's use of diagrams in understanding and explaining groups and subgroups in abstract algebra. In Zandieh, M. (Ed.) *Proceedings of the 11th Annual Conference on Research in Undergraduate Mathematics Education*, San Diego, CA, February 28 - March 2, 2008.

Zimmerman, W. (1991).Visual thinking in calculus. In W. Zimmerman & S. Cunningham (Eds.) *Visualization in teaching and learning mathematics: MAA notes number 19* (pp. 127-137). Washington, DC: Mathematical Association of America.

A HYPOTHETICAL COLLECTIVE PROGRESSION FOR CONCEPTUALIZING MATRICES AS LINEAR TRANSFORMATIONS

Megan Wawro
Virginia Tech

Christine Larson
Vanderbilt University

Michelle Zandieh
Arizona State University

Chris Rasmussen
San Diego State University

In this paper we develop the notion of a hypothetical collective progression (HCP). We offer this construct as an alternative to the construct of hypothetical learning trajectory in order to (a) foreground the mathematical development of the collective rather than that of individuals, and (b) highlight the integral role of the teacher within this development. We offer an abbreviated example of an HCP from introductory linear algebra based on the "Italicizing N" task sequence, in which students work to generate and combine matrices that correspond to geometric transformations specified within the problem context. In particular, we describe the ways in which the HCP supports students in developing and extending local "matrix acting on a vector" views of matrix multiplication (focused on individual mappings of input vectors to output vectors) to more global views in which matrices are conceptualized in terms of how they transform a space in a coordinated way.

Keywords: Linear algebra, collective, learning, progression, linear transformation

The construct of a hypothetical learning trajectory (HLT) was initially developed for and has primarily been used by both teachers and researchers for the purpose of describing individual student learning in particular content domains (e.g., Clements & Sarama, 2004; Simon, 1995; Simon & Tzur, 2004; Steffe, 2004). As reported by Clements and Sarama, the variety of uses and interpretations has expanded to include both individual cognitivist and collective analyses of mathematical development. The abundant use of the term, however, may lead to confusion regarding which unit of mathematical development is under discussion. We hold the view that, in a classroom setting, individual student thinking shapes and is shaped by the development of mathematical meaning at the collective level (Cobb & Yackel, 1996). Choosing to focus on the classroom as the unit of analysis, we adapt the notion of an HLT to articulate a new construct, which we refer to as a hypothetical collective progression (HCP), appropriate to guide the mathematical development at the collective level. In this paper we offer an abbreviated illustration of an HCP in the context of undergraduate linear algebra.

Student difficulties in learning fundamental concepts in linear algebra are well documented (e.g., Carlson, 1993; Dorier, Robert, Robinet, & Rogalski, 2000; Harel, 1989; Hillel, 2000; Sierpinska, 2000). Symbolization of algebraic ideas relies heavily on the use of variables and functions (Arcavi, 1994), and research shows that students at the undergraduate level continue to struggle in their interpretations of variables and functions (Oehrtman, Carlson, & Thompson, 2008; Jacobs & Trigueros, 2008). This difficulty is amplified in the realm of linear algebra, where students must reckon with symbolization in multidimensional contexts.

A result of our work that we present in this piece is an HCP designed to support students in developing their understanding and symbolization of linear transformations defined by matrix multiplication. The main learning goals of this HCP are (a) interpreting a matrix as a mathematical object that transforms input vectors to output vectors, (b) interpreting matrix multiplication as the composition of linear transformations, (c) developing the imagery of an inverse as "undoing" the original transformation**,** and (d) coming to view matrices as objects that geometrically transform a space. These learning goals include a student transition from a localized view wherein matrices are interpreted as transforming one vector at a time to a more global view of a matrix transforming an entire space.

Theoretical Framework and Literature Review

This work draws on three instructional design heuristics of Realistic Mathematics Education (RME) (Freudenthal, 1983; summarized by Cobb, 2011). First, an instructional sequence should be based on experientially real starting points. In other words, tasks that comprise an instructional sequence should be set in a context that is sufficiently meaningful to students that they have a set of experiences through which to meaningfully 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 mathematical learning goals associated with the instructional sequence. 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. In other words, the process of students' reasoning on a task becomes reified so that the outcome of that process of reasoning can serve as a meaningful basis and starting point for students' reasoning on subsequent tasks.

In order to operationalize these RME heuristics into content-specific deliverables that are more explicitly related to instruction, a number of researchers have used the construct of hypothetical learning trajectory. Simon (1995) coined the term to describe the work teachers do in anticipating the path(s) of their students' learning in planning for classroom instruction, and he defined an HLT as consisting of "the learning goal, the learning activities, and the thinking and learning in which the students might engage" (p. 133). In addition to its wide spread use in describing the hypothetical learning of individual students, the construct has been adapted by some to conjecture about the development of mathematical meaning at the collective level (e.g., Cobb, Stephan, McClain, & Gravemeijer, 2001; Gravemeijer, Bowers, & Stephan, 2003; Larson, Zandieh, & Rasmussen, 2008). Indeed, Cobb et al. (2001) describe viewing an HLT as "consisting of conjectures about the collective mathematical development of the classroom community" (p. 117), and Gravemeijer et al. (2003) describe it as a "possible taken-as-shared learning route for the classroom community" (p. 52). We follow this adaptation of the construct for the social perspective, adding here the explicit consideration of the role of the teacher as an integral aspect in the collective sense making that takes place in the classroom enactment.

In order to distinguish this differing collective perspective on HLTs it is necessary to put forth an alternative construct, namely that of a *hypothetical collective progression* (HCP). We define an HCP to be a storyline about teaching and learning that occurs over an extended period of time. The storyline includes four interrelated aspects:

1. Learning goals about student reasoning;

2. Evolution of students' mathematical activity;
3. The role of the teacher; and
4. A sequence of instructional tasks in which students engage.

The second aspect, the evolution of students' mathematical activity, is described in terms of both common difficulties and problematic conceptions that arise, as well as in ways of reasoning that potentially function as if shared. By detailing possible normative ways of reasoning (in the sense of Stephan & Rasmussen, 2002), an HCP emphasizes the potential mathematical development of the collective. It further pays homage to the reflexive relationship between individual thinking and collective development by noting common difficulties and problematic conceptions that arise within students' engagement in mathematical activity.

This construct further differs from than that of an HLT in its explicit inclusion of the role of the teacher as integral in the anticipated progression of mathematical activity in the classroom. The teacher is a unique and essential member of the classroom community with the role of not only pushing forward the mathematical development of the classroom but also fostering productive social and sociomathematical norms within that classroom. Thus, this framing highlights the multi-dimensional structure of classroom activity. As the first and second aspects highlight, a teacher must consider the learning goals she has for her classroom, as well as envision the evolution of students' mathematical development as these goals are actualized. The third and fourth aspects of an HCP—the role of a teacher and the sequence of instructional tasks in which the students engage—speak to how these could be carried out within a given classroom.

Toward Conceptualizing Matrices as Linear Transformations

Research on the learning of linear algebra identifies three common student interpretations of a matrix times a vector: matrix acting on a vector view (MAOV), vector acting on a matrix view (VAOM), and systems views (Larson, 2011). The MAOV view is based on the idea that the matrix acts on or transforms the input vector, thus turning it into the output vector. The VAOM view is based on the idea that the vector acts on the matrix by weighting the column vectors of the matrix, whose sum results in the output vector. A systems view of matrix multiplication is typified by an effort to reinterpret matrix multiplication by thinking of it as corresponding to a system of equations. The HCP we detail offers a means by which instructors can support students in developing and extending the MAOV view of a matrix times a vector to a more global view of how a linear transformation defined by a matrix affects an entire space and how transformations can be composed.

A transformation is a broad mathematical concept that can be represented in a number of ways. For example, a matrix is one specific way in which certain types of transformations (e.g., linear transformations) can be represented. A transformation (function) T from $\mathbf{R}^n$ to $\mathbf{R}^m$ is a rule that assigns to each $\mathbf{x}$ in $\mathbf{R}^n$ a vector $T(\mathbf{x})$ in $\mathbf{R}^m$. A linear transformation $T\colon \mathbf{R}^n \rightarrow \mathbf{R}^m$ is a map that satisfies the following properties: (a) $T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$ for every $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^n$, and (b) $T(a\mathbf{v}) = aT(\mathbf{v})$ for every scalar a and every $\mathbf{v}$ in $\mathbf{R}^n$. It can be shown that every transformation given in terms of matrix multiplication is a linear transformation when defining $T(\mathbf{v})=A\mathbf{v}$ for a given n x m matrix A. For instance, one may consider the transformation T from $\mathbf{R}^2$ to $\mathbf{R}^2$ that rotates the plane ninety degrees counterclockwise; this transformation can be defined by the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. It is this conceptualization, which we refer to as conceptualizing matrices as linear transformations, that is the subject of this paper.

Data Sources and Methods

This research-based HCP grows out of a larger design research project that explores ways of building on students' current ways reasoning to help them develop more formal and conventional ways of reasoning, particularly in linear algebra. The instructional sequence described in this paper was developed and iteratively refined over the course of four semester-long classroom teaching experiments following the methodology described by Cobb (2000) that took place in inquiry-oriented introductory linear algebra classes at public universities in the southwestern United States. We use the term inquiry-oriented in a dual sense, where the term inquiry refers to the activity of the students as well as the teacher (Rasmussen & Kwon, 2007). Students engage in discussions of mathematical ideas, questions, and problems with which they are unfamiliar and do not yet have ways of approaching; thus, evaluating arguments and considering alternative explanations are central aspects of student activity. Teacher activity includes facilitating these discussions, which demands that the teacher constantly inquire into students' thinking. Students in these courses were generally sophomores or juniors in college, majoring in math, engineering, or computer science, and were required to have successfully completed two semesters of calculus prior to enrollment in the course.

During each classroom teaching experiment, we videotaped every class period using 3-4 video cameras that focused on both whole class discussion and small group work. We also collected student written work from each class day. As a research team, we met approximately three times a week in order to debrief after class, discuss impressions of student work and mathematical development, and plan the following class. We also used these meetings retrospectively to inform decisions regarding the following iteration of the classroom teaching experiment, as what we analyzed one semester became refined and informed the next iteration of the curriculum. One of our goals was to produce an empirically grounded instructional theory, and doing so involves a number of stages. Over the four years, we have refined not only our instructional tasks, but we also have deepened what we know about student thinking in linear algebra, refined the learning goals for our course, and increased our awareness of the role of the teacher. One of the results from this extensive iterative work is the notion of an HCP. Examples presented in this paper were taken from the fourth and latest classroom teaching experiment.

The HCP presented in this paper is the result of a retrospective analysis of the development of the instructional sequence and the associated set of learning goals, as well as the way in which the instructor used this instructional sequence to support students' mathematical activity in its classroom enactment. Instructor and student notes were used to reconstruct the broad progression of classroom activity across the set of tasks; these were used to identify relevant segments of classroom video from the classroom teaching experiment. Our research team generated memos to document students' mathematical activity and the role of the teacher in progressing through this particular enactment of the instructional sequence, paying particular attention to the variety of student interpretations elicited by the task, issues that were problematic for students, and the role of the instructor in negotiating sense-making around student generated notation and connecting to more conventional notation used by the broader mathematical community.

Prior to the instructional sequence driven by this HCP, the class had engaged in an RME-inspired instructional sequence focused on helping students develop a conceptual understanding of linear combinations, span, and linear independence (Wawro, Rasmussen, Zandieh, Sweeney, & Larson, in press). The class also spent time developing solution techniques for linear systems to help answer questions regarding span and linear independence of sets of vectors. This led to the definition and exploration of equivalent systems, elementary row operations, matrices as an array of column vectors, augmented matrices, Gaussian elimination, row-reduced echelon form, pivots, and existence and uniqueness of solutions. This broad set of ideas was unified by developing and proving conjectures regarding how these ideas fit together for both square and non-square matrices.

Results

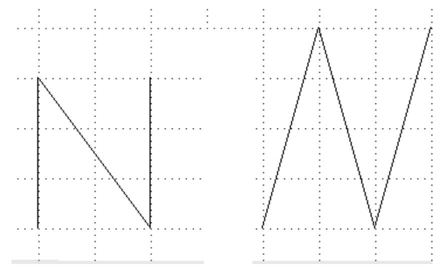
The HCP developed in this report encompasses four learning goals: (a) Interpreting a matrix as a mathematical object that transform input vectors to output vectors; (b) Interpreting matrix multiplication as the composition of linear transformations; (c) Developing the idea of an inverse as "undoing" the original transformation; and (d) Coming to view matrices as objects that geometrically transform a space. These learning goals are not intended to be achieved sequentially. Rather, these four learning goals interweave and support students in developing a robust conceptual understanding of matrices as linear transformations. For instance, one may see learning goal (a) as a local view of linear transformation, whereas learning goal (d) may be interpreted as a more global view. The global view is not meant to replace the local view; rather, it elaborates it. We want students to be able to draw on and coordinate both views, moving flexibly between them as need be. In fact, coordination of local and global views is an aspect of students' mathematical activity that cuts across all tasks in the instructional sequence in a way that we argue is crucial to the development of productive normative ways of reasoning.

Our construct of HCP has four components, and we organize the results section in terms of the fourth: the sequence of instructional tasks in which students engage. We choose to do this because it is the aspect of the HCP in which the students do engage sequentially, so this allows for a more natural parallel to how the HCP may unfold in an actual enactment. For each task, we discuss what the students are being asked to do and how this is significant in terms of our learning goals. Using data from the fourth classroom teaching experiment in linear algebra, we describe students' mathematical activity as they engage in each task, sources of difficulty, and the role of the teacher in supporting students to work through these difficulties. We focus especially on the role of the teacher in negotiating the use of mathematical notation to support sense making and to connect to symbolic and definitional conventions of the mathematical community.

Following an introduction to transformation view of $A\mathbf{x} = \mathbf{b}$, this particular HPC regarding linear transformations has three main tasks: (a) the Italicizing N task; (b) the Pat and Jamie task; and (c) the Getting Back to the N task.

Introduction to a Transformation View of $A\mathbf{x} = \mathbf{b}$

The first learning goal of this HCP is conceptualizing matrices as mathematical objects that transform input vectors to output vectors. That is, in contrast to interpreting $A\mathbf{x} = \mathbf{b}$ in terms of a vector equation or a system of equations, the goal is to encourage conceiving of $A\mathbf{x} = \mathbf{b}$ as a matrix A acting on the vector $\mathbf{x}$ to produce the vector $\mathbf{b}$. This goal involves a major interpretive shift for students, but their prior experiences working with functions serve as a good starting point for this new conceptualization of matrices. One way in which the teacher can support this shift is by introducing terminology that will support students' work in the upcoming sequence of tasks by helping them analogize their work with matrices to their prior knowledge of functions. For instance, the teacher may introduce the terms like transformation, domain, and codomain, and discuss how $A\mathbf{x} = \mathbf{b}$ can be interpreted as an example of a transformation by defining $T(\mathbf{x}) = A\mathbf{x}$. These introductory whole-class discussions offer students the opportunity to begin to lay a foundation for thinking of input-output pairs of vectors that are related through a matrix transformation. Rather than provide further specifics of this introductory aspect, we shift out focus to the main task of this HCP, students' mathematical development through interaction with this task, and the role of the teacher.

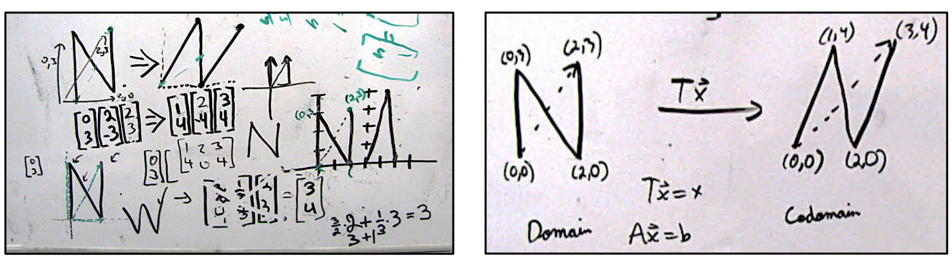


Suppose the "N" on the left is written in regular 12-point font. Find a matrix A that will transform N into the letter on the right, which is written in italics in 16-point font.

Figure 1. The Italicizing N Task

The Italicizing N Task

The Italicizing N task (see Figure 1) is the first task in our HCP, and it is through this task that students embark on their initial exploration of matrices as linear transformations. In this task, the students' goal is to determine a matrix A that represents the requested transformation of the N described in the problem statement. The teacher plays a crucial role in setting up this task by supporting students in developing a common interpretation of the setting and goals of the task, as well as in interpreting matrix multiplication as a transformation. Specifically, assumptions about the context and aspects of the mathematical goals need to be negotiated (e.g., how to represent the N mathematically in each image, as well as how to determine a matrix that maps the image on the left to the one on the right).



(a) (b)

Figure 2. Student work on the Italicizing N task.

It is nontrivial for students to determine that both the inputs and outputs for the transformation lie in $\mathbf{R}^2$ and that the mapping could be represented by a 2 x 2 matrix. Furthermore, students grapple with how to interpret and symbolize the representations of the N. Examination of past student work has revealed two common strategies: using vectors in $\mathbf{R}^2$ or using points in the x-y plane. For example, within the student work shown in Figure 2a, the N is represented with vectors whose tip and tail lay on the letter with tips originating from the same point on the letter (corresponding to a fixed origin). Other students represent the N with vectors whose tip and tail lay on the letter but with tips originating from different points on the letter (corresponding to a “floating” origin). On the other hand, within the student work shown in Figure 2b, locations on the N are labeled as points on the *x-y* plane with an origin anchored at the lower left vertex of each N.

Regardless of the way in which students represent the letter N, a potential normative way of reasoning is setting up a system of matrix equations – one matrix equation for each input-output pair – in order to determine the component values of *A*. An example of this approach is shown in Figure 3. The instructor is able to use the variety of representations students generate for the letter N as a starting point for a class discussion about the relationships among choices of representation (points versus vectors), the significance of where one chooses to place the origin, and whether those choices affect the matrix that transforms the letter in the desired way. This allows the teacher to push students to make connections among various approaches and bring out key mathematical ideas. The teacher, as a member of the mathematical community, is in a position to raise questions, such as why anchoring the origin would be advantageous, that the students are not necessarily in the position to make on their own. It is this interaction between the role of the teacher and students’ mathematical progress on an instructional task that helps promote a climate of sense making, fosters social norms such as listening to others’ reasoning and providing explanation of your own, and moves forward the mathematical goals.

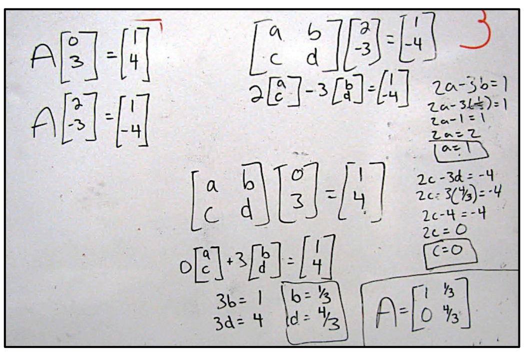


Figure 3. Students set up matrix equations to solve for the values of matrix *A*.

Other important aspects of this task that are not immediately obvious to students include how to select input-output pairs, how many input-output pairs are needed to determine the matrix, and whether the matrix will be unique. As students share their approaches for finding the matrix *A*, the teacher has the opportunity to ask students about these aspects. For instance, the teacher can facilitate a discussion about whether *A* is a unique matrix representation of the transformation (according to the standard basis, which has remained implicit at this point). Often, at least one group of students selects a linearly dependent set of vectors to generate their system of matrix equations – and these students often argue that *A* is not unique. With the instructor’s facilitation, this disagreement over *A* provides an opportunity for students to discuss what it means for *A* to be unique, or under what criteria is *A* unique, and if so, what the criteria are for selecting sets of input-output pairs that uniquely determine *A*.

In our classroom teaching experiments, we follow the Italicizing N Task with activities that ask students to investigate other transformations of the plane, such as stretching, rotating, etc. The emphasis here begins to shift away from only considering particular input-output pairs to how the transformation defined by *A* affects the entire plane, without needing to go through the motions of plotting particular pairs. While still working in $\mathbf{R}^2$, the teacher suggests other transformations (such as stretching and skewing images in Quadrant 3) to develop a connection to geometric interpretations of the standard 2 x 2 transformation matrices. This leads into and is not disjoint from the fourth learning goal of coming to view matrices as objects that geometrically transform a space.

The Pat and Jamie Task

The Pat and Jamie task (see Figure 4) was the first task introduced associated with the learning goal of interpreting matrix multiplication as a composition of linear transformations. This is a follow-up to the Italicizing N Task, and it sets up a scenario in which the students must first decide if the approach of two "fellow students," Pat and Jamie, is valid, and then determine the matrices that represent the transformation via their approach. Note that in the problem statement, the order in which Pat and Jamie transform the N is not vague (they make it taller first and then italicize it), but the way in which to computationally accomplish this is purposefully left vague. Students are meant to struggle with how to combine and symbolize one transformation followed by another and why that is sensible.

Last semester, two linear algebra students—Pat and Jamie—described their approach to the Italicizing N Task in the following way:

> "In order to find the matrix *A*, we are going to find a matrix that makes the "N" taller (from 12-point to 16-point), find a matrix that italicizes the taller "N," and the combination of those will be the desired matrix *A*."

1. Does their approach seem sensible to you? Why or who not?
2. Do you think their approach allowed them to find a matrix *A*? If so, do you think it was the same matrix *A* we found this semester?
3. Try Pat and Jamie's approach. You should either: a) come up with a matrix *A* by using their approach, or b) be able to explicitly explain why this approach does not work.

Figure 4. The Pat and Jamie Task

The Pat and Jamie task sets the stage for a shift in students' mathematical activities and goals; students are asked to *combine* matrices that define transformations in addition to determining what those transformations matrices are. This is a shift from the goal of constructing a single transformation matrix based on inputs and outputs, such as in the Italicizing N task. This shift is significant because matrices are beginning to be positioned as objects of students' mathematical activity rather than solely the result of a mathematical process (Sfard, 1991). Students often are successful in constructing matrices for the individual transformations (which is a natural continuation of their work on the Italicizing N task) but struggle more with what a sensible way to "combine" these matrices would be. For instance, Figure 5 shows a student's correct matrix *A* for making the N taller, but the student's matrix (also labeled *A*) for italicizing the taller N is incorrect (but rather would italicize the shorter, original N correctly). Thus, this student did not attend to the chain of transformations, in which the output for one transformation (making taller) serves as the input for the following one (italicizing). Furthermore, the student writes, "How do we combine these?" on his/her paper, which further indicates the student's struggle with this problem. The teacher needs to be aware of common problematic conceptions such as this, and because of that, it is the role of the teacher to ask questions, make comments, and use notation that direct the class towards to the desired mathematical progress.

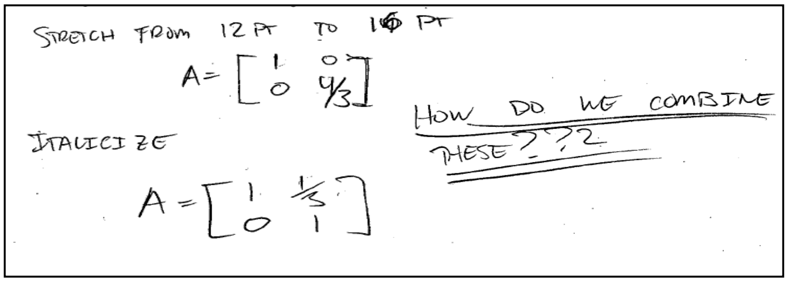


Figure 5. Student expresses uncertainty of how to combine the separate transformations to first change from 12 to 16-pt font and then italicize.

The class begins the Pat and Jamie task by positing that this approach should give the same final matrix as in the initial Italicizing N task. Thus, students often experiment with ways of creating and combining the two intermediary matrices until the combination yielded the desired matrix *A* (which they had found in the previous activity), mostly trying addition or multiplication. However, students often have difficulty knowing why their operation choice and order is logical. For instance, after working for a while, every small group may have the correct matrix to make the 12-pt N the taller 16-pt N. The class negotiates to name this matrix *B* because *T* is used for other things (namely, to refer to generic linear transformations). However, students struggle to find the matrix that would yield the desired 'lean.' The teacher plays a role in working through this struggle in whole class discussion by having various students explain their thoughts and approaches. This is peppered throughout with the teacher asking clarifying questions and revoicing the students' approaches in both words and symbols. For instance, some groups may (correctly) use the vectors from the middle N as the inputs for the lean transformation to correctly determine *L*. The teacher may choose to summarize and revoice this type of explanation with mathematical symbolism on the board by restating the explanation and explicitly discussing how the output of the first transformation becomes the input for the second and illustrate this with particular input-output pairs (see Figure 6). This provides a way to move the mathematical agenda forward but still allow students to reason for themselves why the correct order of matrix multiplication in sensible according to Pat and Jamie's approach. This also promotes the sociomathematical norms of developing mathematical justifications for computational choices and explaining them.

To summarize, two main choices often surface through the students' work on the task. First, many groups determine, at least initially, that the matrix for the "lean" transformation is □=*11/301* (this is consistent with the matrices in Figure 5). Students who remain with this (incorrect) choice discover that *BL* = *A*. Other groups determine (as described above), that the "lean" transformation is □=*11/401* and that the (correct) matrix multiplication of *LB* yields the desired matrix *A*. Note that the two approaches have the same matrix *B* but two different matrices for *L*. The teacher has the opportunity to write both matrix equations on the board, highlight how students got the correct *A* in two different ways, and ask which approach is what Pat and Jamie did and how they could be certain.

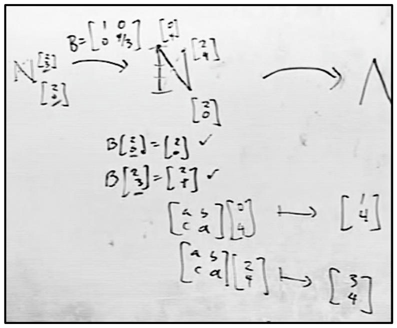
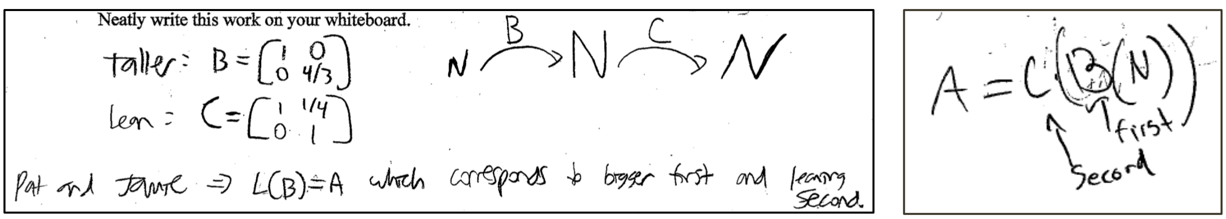


Figure 6. The teacher's symbolic revoicing of students' approaches to the Pat & Jamie Task.

We highlight this episode as significant because it illustrates (a) probable student difficulties with developing an intuitive notion of function composition in the context of linear transformations, and (b) the role of the teacher in connecting to student thinking as she moves her mathematical agenda forward. We posit that the teacher's purposeful use of symbolic notation on the whiteboard during whole class discussion serves as an example of the pedagogical content tool of transformational record (Rasmussen & Marrongelle, 2006). She recorded student thinking regarding input-output pairs for the various transformations and added notation—such as the three N's, the arrows and corresponding transformation matrices between the N's, and the vector and matrix equation representations of the input-output pairs—that later served as tools in students' reasoning about which order of matrix multiplication correctly matched Pat and Jamie's approach.



(a) (b)

Figure 7. A snapshot of one student's work to explain which approach is correct and why.

For instance, the images in Figure 7a and 7b serve as examples of common ways of reasoning after a discussion such as the one highlighted in Figure 6. Figure 7a highlights how a student found the two transformations and the correct order of multiplication; the student also wrote "$CB = A$ which corresponds to bigger first and leaning second." The student work in 7b highlights a bit of the compositional nature of the task, where the "$(B(N))$" seems similar to the notion of "$f(x)$," with the matrix C then acting on the output of $B(N)$. Note that both of these responses only begin to hint at the compositional nature of the two transformations, with the output of B becoming the input for C. It is the role of the teacher to facilitate discussion of this. These two examples of student work also illustrate the potential to shift away from only focusing on how individual vectors are transformed to how a space is transformed. As such, this task, in addition to developing the notion of matrix multiplication as the composition of functions, further fosters the first and fourth learning goals of the local and global aspects of matrices as transformations.

Finally, it should be mentioned how the Pat and Jamie task is built upon in subsequent classroom discussion. Along the way in determining which matrix equation describes Pat and

Jamie's approach, the students work to determine that, in this case, the order in which the matrices are multiplied affect the answer. The teacher has the opportunity and responsibility, as a member of the mathematics community, to introduce and connect the term *commutativity* to the students' work. The teacher also works to connect this to the notion of composition of functions as an interpretation of matrix multiplication. Again, as a member of the mathematics community, the teacher serves as a broker (Rasmussen, Zandieh, & Wawro, 2009) between students' authentic mathematical activity and the terminology and notation commonly used in the mathematics community.

The Getting Back to N Task

The last main task in the instructional sequence within this HCP is the Getting Back to N task (see Figure 8), and it is mainly associated with the learning goal of developing the idea of an inverse as "undoing" the original transformation. It is also intended to further the learning goal of reasoning about matrix multiplication as a composition of linear transformations. This task asks students to determine a matrix C that will transform the letter on the right back into the letter on the left; that is, from the 16-pt italicized N to the original N. Furthermore, this task asks students to determine the matrix C in two ways: through one direct transformation and through Pat and Jamie's method (i.e., in two steps). To the expert, this task introduces the notion of inverse transformations. This naturally follows from students' work on the previous tasks, although it is by no means trivial for students, because investigating the Pat and Jamie approach to "getting back to the N" reiterates their work regarding function composition and matrix multiplication. A rationale behind our development of this task is consistent with Oehrtman et al.'s (2008) observation that more sophisticated, process-oriented views of an inverse function coincide with conceiving of it as the function that undoes the action of the original function, versus a less sophisticated, action-oriented view that associates the concept of inverse function with a surface action such as determining the associated matrix inverse via a memorized formula. It additionally requires students to consider the importance of order in function composition and matrix multiplication when one or more of the functions under consideration has the action of "undoing" the action of a previous transformation.

Regarding the Italicizing N Task, complete the following:

Find a matrix C the will transform the letter on the right back into the letter on the left.

1. Find C using either your method or one of your classmate's method for finding A
2. Find C using Pat and Jamie's method for finding A.

Figure 8. The Getting Back to N task.

Initial student work on the first prompt of finding the matrix C in one step is often unproblematic. Given the students' engagement with the previous two tasks, a potential normative way of reasoning is to determine C by creating a matrix equation that coordinates particular input-output pairs from the tall, italicized N to the original N. Part of the role of the teacher, however, is to call attention to the relationship between these input-output pairs and those from the original Italicizing N task; namely, that the inputs in the Getting Back to N task served as the outputs in the Italicizing N task and vice versa. This emphasis on the role of the various parts of matrix equations such as $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{b} = \mathbf{x}$ lays the groundwork for subsequently labeling the matrix *C* as the inverse of matrix *A*. Given that students are not, at this stage, formally aware of *C* as the inverse of *A*, the symbol used to notate this relationship is tied to students' ways of thinking and symbolizing. For instance, based on students' normative way of reasoning that the transformation defined by *C* "undoes" the effect of the transformation defined by *A*, the teacher leads the class in defining $C = U_A$, where U_A stands for "undoes *A*."

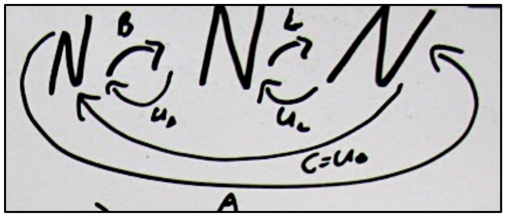
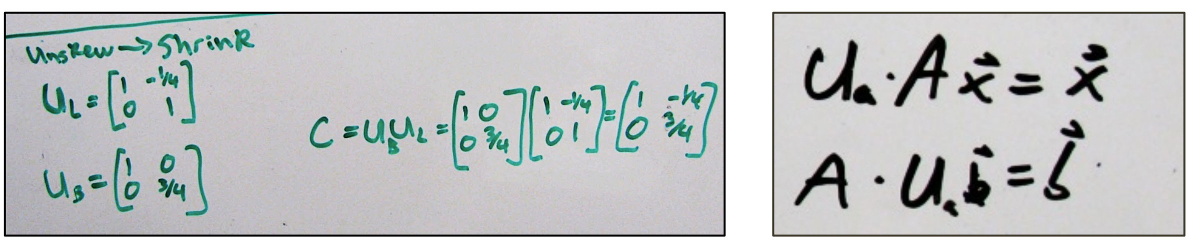


Figure 9. Negotiation of meaning for the matrices U_B, U_L, and U_A, which are the matrices that "undo" the actions of *B*, *L*, and *A*, respectively.

When working on the second prompt of determining the matrix *C* via Pat and Jamie's method, the students are faced again with not only again determining the matrix transformations for the constituent parts but also how to combine those matrices in a sensible manner. As seen in Figure 9, the class's work often begins by again creating names for each matrix associated with a transformation. Here student work illustrates a naming scheme developed during the class; students chose to use the letter B to label the matrix that made the matrix bigger, L for the matrix that made the N lean, and *A* is the original transformation matrix from the Italicizing N task (note, however, that what was labeled *C* in Figure 8 is labeled *L* in Figure 10). Also, the arrows going the opposite directions correspond to the transformations U_L, U_B, and U_A, which "undo" the original transformations defined by *L*, *B*, and *A*, respectively. Also note that the notation in Figure 9 is layered upon the symbolic representations developed in the original Pat and Jamie task (see Figures 6 and 7a), but that the standard symbolism for inverses is not used here. The teacher can use this development of notation to further foster sociomathematical norms of explicitly defining symbolic notation and providing justification for these choices. This also serves to connect to students' current ways of reasoning that the teacher could then leverage into formal notation and definitions used by the mathematics community.



(a) (b)

Figure 10. Student work on the Getting Back to N task.

The notation developed in Figure 9 can subsequently be used by students to calculate the numerical values of U_L, U_B, and U_A. The example of student work in Figure 10a illustrates a potential normative way of reasoning that matrix multiplication for a composition of functions is constructed right to left, with the matrix for the first transformation being multiplied on the left by matrices for the subsequent transformations. The example of student work shows the correct order of computation to find the inverses for Pat and Jamie's approach. The student wrote, "unskew → shrink" to indicate that the first action is to undo the lean and the second is to undo making it bigger (notated by the matrices U_L and U_B, respectively). The second example of student work (see Figure 10b) shows a student's notation illustrating that A composed with its "undoing" matrix U_A in either order leaves the input vector unchanged. This connects back to the matrix equation $A\mathbf{x} = \mathbf{b}$ interpretation of A acting on the vector $\mathbf{x}$ to produce the vector $\mathbf{b}$. In the first line of Figure 10b, matrix A acts on the vector $\mathbf{x}$, and matrix U_A acts on the resultant vector $A\mathbf{x}$ to produce vector $\mathbf{x}$. Similarly, the second line of Figure 10b connects to the aforementioned matrix equation $C\mathbf{b} = \mathbf{x}$ in that the matrix U_A acts on the vector $\mathbf{b}$, and matrix A acts on the resultant vector $U_A\mathbf{b}$ to produce vector $\mathbf{b}$. This sophisticated way of symbolizing gives the teacher a launching point from student reasoning within the task setting to connect to the definitions and notation of the mathematics community. For instance, she can facilitate a conversation about interpreting the matrix transformations (U_AA) and (AU_A) in Figure 10b as transformations that have the action of "doing nothing" to any given input vector, leading to a symbolic notation of these composition transformations as a "do nothing" transformation, notated by the letter I; that is, $(U_AA)\mathbf{x} = I\mathbf{x} = \mathbf{x}$ and $(AU_A)\mathbf{x} = I\mathbf{x} = \mathbf{x}$. Finally, as a member of the classroom community and the mathematics community, the teacher acts as a broker between them by connecting the class's work with the formal definitions of inverse for both linear transformation (i.e., a linear transformation $T:\mathbb{R}^n \rightarrow \mathbb{R}^n$ is invertible given that there exists a transformation $S:\mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $T(S(\mathbf{x}) = \mathbf{x}$ and $S(T(\mathbf{x})) = \mathbf{x}$) and matrix multiplication (i.e., a matrix A is invertible given that there exists an n x n matric C such that $AC = I$ and $CA = I$).

Conclusion

This abbreviated example of an HCP highlights several aspects of the interrelationships among the components in our definition of the construct. First, we highlight the fact that the learning goals are not sequential in nature, and we posit that the way in which these goals cut across tasks is important for supporting the development of productive normative ways of reasoning. For example, students repeatedly engage in the mathematical activity of coordinating input-output pairs to construct matrices, and this comes to function as the basis for later reasoning about how to conceptualize those matrices as mappings that can be combined (composed) and undone (inverted). Second, we highlight the fact that the variety of student approaches and sources of student difficulty function as a source of need for public negotiation of meaning – and that these conversations can and should contribute to the development of ideas that come to function as-if-shared in the classroom. This type of public negotiation of meaning is inextricably linked to the work the teacher does in using students' mathematical activity as a basis for group sense-making that moves forward the mathematical agenda. Finally, we highlighted the complexity of the role the teacher in negotiating meaning around student generated notation and in introducing more conventional notation in a way that honors student generated notation and connects to the broader mathematical community.

Looking back across the HCP, we see the first and fourth learning goals (interpreting a matrix as a mathematical object that transforms input vectors to output vectors, and coming to view matrices as objects that transform a space) cutting through the sequence of activities. More specifically, students repeatedly coordinate local and global views of matrix multiplication – including in the context of composing and inverting mappings. In addition, by repeatedly restructuring the original problem context in a variety of ways (as a two-step map for composing, and as a backwards map for inverting), students' activity shifts from an initial focusing solely on coordination of inputs and outputs to construct matrix representations for particular mappings, to a later focus on coordinating and combining those mappings (which also involves coordination of inputs and outputs but is in service of the goal of combining mappings).

We conclude by noting that this is consistent with Oehrtman, Carlson, and Thompson's (2008) recommendations for teaching ideas about functions. Namely, they recommend explicitly asking students to explain ideas about functions in terms of inputs and outputs, as well as asking students to explain function behavior on entire intervals (as opposed to just asking about function behavior in a pointwise fashion). Our HCP encourages students to coordinate inputs and outputs for the purpose of constructing matrices that yield desires mappings. In order to construct mappings that yield desired geometric transformations, students have a need to conceptualize the domain and codomain in a coordinated way for the purpose of selecting (linearly independent) sets of inputs and outputs in the domain and codomain, respectively.

Acknowledgments

We thank our colleagues who attended our presentation at the 2012 RUME Conference and gave insightful feedback. Your comments and questions helped shape and refine this paper.

This material is based upon work supported by the National Science Foundation under grants no. DRL 0634099 and DRL 0634074. 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

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the Learning of Mathematics, 14*(3), 24-35.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*, 81-89.

Carlson, D. (1993). Teaching Linear Algebra: Must the Fog Always Roll In? *The College Mathematics Journal, 24*, 29-40.

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.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10,* 113-163.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*, 175-190.

Dorier, J.-L., Robert, A., Robinet, J., & Rogalski, M. (2000). The obstacle of formalism in linear algebra: A variety of studies from 1987 until 1995. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 85-124). Dordrecht: Kluwer Academic Publishers.

Freudenthal, H. (1983). Didactical phenomenology of mathematical structures. Dordrecht: Reidel.

Gravemeijer, K., Bowers, J., & Stephan, M. (2003). A hypothetical learning trajectory in measurement and flexible arithmetic. *Journal for Research in Mathematics Education. Monograph, Vol. 12,* 51-66.

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,* 139-148.
Hillel, J. (2000). Modes of description and the problem of representation in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 191-207). Dordrecht: Kluwer Academic Publishers.
Jacobs, S., & Trigueros, M. (2008). On developing a rich conception of variable. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 3-13). Washington, DC: Mathematical Association of America.
Larson, C. (2011). *Matrix multiplication: A framework for student thinking.* Manuscript submitted for publication.
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 Research in Undergraduate Mathematics Education Conference. San Diego, CA.
Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' understandings of function. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 150-171). Washington, DC: Mathematical Association of America.
Rasmussen, C., & Kwon, O. (2007). An inquiry oriented approach to undergraduate mathematics. *Journal of Mathematical Behavior, 26*, 189-194.
Rasmussen, C., & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics into instruction. *Journal for Research in Mathematics Education, 37*, 388-420.
Rasmussen, C., Zandieh, M., & Wawro, M. (2009). How do you know which way the arrows go? The emergence and brokering of a classroom mathematics practice. In W.-M. Roth (Ed.), *Mathematical representation at the interface of body and culture* (pp. 171-218). Charlotte, NC: Information Age Publishing.
Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Education Studies in Mathematics, 22*, 1–36.
Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 209-246). Dordrecht: Kluwer Academic Publishers.
Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26,* 114-145.
Simon, M. A., & Tzur, R. (2004). Explicating the role of the mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning, 6*, 91-104.
Steffe, L. (2004). On the construction of learning trajectories of children: The case of commensurate fraction. *Mathematical Thinking and Learning, 6,* 129-162.
Stephan, M., & Rasmussen, C. (2002). Classroom mathematical practices in differential equations. *Journal of Mathematical Behavior, 21,* 459-490.
Wawro, M., Rasmussen, C., Zandieh, M, Sweeney, G. F., & Larson, C. (in press). *An inquiry-oriented approach to span and linear independence: The case of the Magic Carpet Ride sequence.* PRIMUS.

DO GENERIC PROOFS IMPROVE PROOF COMPREHENSION?

Keith Weber[1], Juan Pablo Mejia-Ramos[1], Evan Fuller[2], Kristen Lew[1], Philip Benjamin[1], Aron Samkoff[1]
Rutgers University[1], Montclair State University[2]

In an influential article, Rowland (2001) suggested "generic proofs" might improve students understanding and appreciation of the proofs that they study. In this paper, we present a qualitative and quantitative study exploring how well students understand generic proofs. In a qualitative study, we found students generally have positive opinions about generic proofs, believing generic proofs can be a useful tool for improving understanding. In a larger quantitative study, we conducted a randomized experiment where we assessed how well undergraduates understood the same proof when presented traditionally and generically. Those who read the proof generically performed somewhat (although not statistically reliably) better on questions applying the ideas of the proof to specific examples but statistically reliably worse on other types of assessment questions.

Key words: proof; proof comprehension; generic proofs; undergraduate mathematics education

Introduction

Proof Presentation and Comprehension in Advanced Mathematics Courses

In advanced undergraduate mathematics lectures, professors typically spend a large amount of time presenting proofs of theorems to their students (Weber, 2004). Mills (2011) reported that roughly half of the lecture time she observed of three mathematics professors was spent on proof presentation. Presumably an assumption behind this pedagogical practice is that students can learn mathematics from observing and studying these proofs. However, many question whether this assumption is reasonable, arguing that students often find proofs to be generally confusing or pointless (e.g., Harel, 1998; Porteous, 1986; Rowland, 2001) argue. There is also evidence that undergraduate mathematics students cannot distinguish between a valid proof and an invalid argument (Selden & Selden, 2003; Weber, 2010); if students cannot determine if proofs are correct, it is doubtful they understand them all that well.

Some mathematics education researchers ascribe students' difficulties in understanding proofs to the formal and linear style in which proofs are written (e.g., Thurston, 1994; Rowland, 2001). Formal and linear proofs are claimed to hinder student proof comprehension for a number of reasons. Some researchers remark that the formal syntax and technical jargon used in proofs can be intimidating to students (Davis & Hersh, 1981; Hersh, 1993; Kline, 1973; Thurston, 1994). Others claim the linear nature of proof can hinder student ability to see the higher-level ideas of the proof, making the assertions of the proof mysterious and unmotivated (Anderson, Boyle, & Yost, 1986; Davis & Hersh, 1981; Leron, 1983).

Several mathematics education researchers have proposed alternative proof formats such as e-proofs (Alcock, 2009), explanatory proofs (e.g., Hanna, 1990; Hersh, 1993), and structured proofs (Leron, 1983). However, there are few empirical studies on these alternative proof formats. Further, systematic studies that have assessed the utility of these alternative proof formats with large samples have failed to find significant learning gains (Roy et al., 2010; Fuller et al., 2011).

Generic Proofs

Rowland (2001) suggested the use of "generic proofs" to improve student understanding. A generic proof is an argument that shows why a general claim is true for a specific example; however the reasoning applied to that example can be applied to any other example as well. Consequently the reader can infer that the general claim will hold for all examples. Rowland argued that a generic proof should have the following five elements:

- If the theorem being proven is of the form, "for every n, n has property P", the generic proof should begin with a particular n_0.
- The particular example, n_0, should be neither too trivial nor too complicated,
- Steps of reasoning are not rooted in the mathematical objects, n_0, themselves, but in the properties of such objects,
- The reasoning should be constructive, and
- For novice students, some scaffolding is needed to ensure that students perceive the argument and use the notation needed to communicate them.

Based on an informal survey of his own students, Rowland found that students valued generic proofs as a tool for comprehension. In a subsequent quantitative study, Malek and Movshovitz-Hadar (2011) investigated the efficacy of generic proofs. They provide a list of four criteria for the proofs to be included in their study:

- Students were unlikely to be able to construct a proof of the theorem on their own,
- There exists a proof for the theorem for which a generic proof could be constructed,
- The proof was not very long so that an interview could be an appropriate length, and
- There exists another theorem that has a somewhat similar proof.

Malek and Movshovitz-Hadar found that student reading generic proofs in linear algebra outperformed students that read analogous formal proofs across a range of comprehension tasks when the proof involved non-routine techniques. However, as their study only involved ten students, with only three or four of them having exposure to each of the generic proofs in the study, the generality of these results is limited.

Research Questions

Our study builds on the existing literature by further investigating the opinions and performance of a larger number of mathematics majors who see generic proofs. In particular, we address two questions:

- Do students believe generic proofs can increase understanding?
- Does reading a generic proof actually improve student's comprehension of this proof?

Before proceeding further, we emphasize that whether generic proofs improve comprehension is a complex question that cannot be addressed with a single study. It is crucial to note that the answer to this question depends on how the researcher defines comprehension, how generic proofs are introduced to students, and which generic proofs are given to students. Consequently the data that we present are not intended to measure the pedagogical value of generic proofs, but instead are a first attempt to generate hypotheses about the strengths and limitations of this proof format.

To address these questions, we examine students' understandings of and reactions to generic proofs immediately after reading them. This methodology of assessing students shortly after reading a proof is consistent with previous studies in the literature on proof comprehension (e.g., Fuller et al, 2011; Malek & Movshovitz-Hadar, 2011; Martini et al., 2012; Roy, et al., 2010). However it is important to note that if generic proofs were used in a different way—such as by having the professor present a generic proof orally or asking the students to study the proofs overnight—different learning benefits may have been observed. We also note that, like Malek and Movshovits-Hadar (2011), we used generic proofs in lieu of traditional proofs. Rowland (2001) recommended this as a possibility, but also suggested presenting a generic proofs prior to a traditional proof. This could have yielded different learning outcomes than what we assessed.

Theoretical Perspective

How a researcher determines whether a proof format improves comprehension is based on how comprehension is defined. In this paper, we base our understanding of proof comprehension on the model of Mejia-Ramos et al (2012). Based on a survey of the mathematics education literature and interviews with mathematicians, Mejia-Ramos and colleagues developed seven dimensions to assess students understanding of proof: (a) the meaning of terms and statements within a proof, (b) how individual claims within the proof were justified, (c) what assumptions and conclusions were used in the proof and how they related to the proof framework, (d) what would constitute a high-level summary of the ideas of , (e) how the ideas in the proof could be applied to prove other statements, (f) what parts of the proof could be partitioned into modular components, and (g) how the ideas of the proof related to specific examples. These seven dimensions are described in detail, along with a rationale for how they were chosen, in Mejia-Ramos, et al. (2012).

Generic proofs seem to be theoretically designed to help with the example and transfer aspects of this model since the proof implements a generic example and the constructive nature illuminates the method being applied (Malek & Movshovitz-Hadar, 2011). Consequently we included example and transfer questions in our study. We also included summary and justification questions in our assessment, based on our opinion of their importance for comprehension.

Qualitative Study

Rationale

The qualitative study was a study in which students were interviewed and videotaped while reading generic proofs, completing an assessment of their comprehension, and giving their feedback on the format of the proof. This study allowed us to investigate students' opinions on the generic proof format. We included assessment items for the purpose of designing better assessment questions for the quantitative study.

Methods

Participants. Students in the study were from a large northeastern state university in the United States. Ten students were recruited to participate in the study and were paid for their participation. Students were recruited in the Spring 2011 and Summer 2011 semesters and were in the fourth or fifth (final) year of a joint B.A. and Ed.M. mathematics education program. All students in this program were mathematics majors.

Materials. The generic proof used in this study was of the theorem: There are 2^{n-1} ways to express n as an ordered sum of positive integers. After the assessment, a linear proof of the same theorem was provided for comparison when discussing the generic format. Participants were also provided written instructions on generic proofs prior to reading the generic proof. These instructions and proofs are given in the Appendix at the end of this paper. Following Rowland (2001), this proof was couched in the domain of number theory. The proof also was written to meet the criteria of Rowland (2001) and Malek and Movshovitz-Hadar (2011). Finally, this theorem was chosen because it is accessible to math majors without having taken a course in number theory.

Procedure. Each participant met individually with an author of this paper for a video-recorded semi-structured interview. Participants were given the generic proof and asked to read the proof until they felt that they understood it. After reading the proof, participants were asked to report: (1) on a scale of 1 through 5, how well they felt they understood the proof (2) on a scale of 1 through 5, how convincing they found the argument, and (3) whether they thought the method of the argument could be generalized to any integer n. After completing all of the questions, the interviewer returned the generic proof and asked the participants:

- What do you think about the format in which this proof was presented?
- (After providing a linear proof of the same theorem,) Here is a more traditional version of the same theorem. If you were in a class, would you prefer the traditional version or the proof you just read?
- Was there anything better about this new version of the proof that might have helped you understand the proof better?
- Was there anything about this new version of the proof that might have made it more difficult to understand?

Interviews ranged from 25 to 60 minutes.

Analysis. Interviews were coded using an open coding scheme (Strauss & Corbin, 1990). Two authors made an initial pass through the data, noting every time a participant commented on a perceived attribute or deficiency of the generic proof format. This initial analysis yielded six preliminary categories that we discuss in the results section. Once these categories were identified, the data was analyzed again more carefully, flagging for every instance in which a participant made a comment of this type.

Results

Table 1 summarizes participants' comments about the generic proofs. As Table 1 illustrates, the participants' comments were generally positive but participants were critical of the generic proof format. The student feedback suggests that students value the reduced abstraction and elimination of notation and jargon, as well as note how generic proofs can aid student comprehension. On the other hand, we also found that students had some criticisms about the generic proof.

	Number of Participants	**Specific Comments**	**Number of Participants**
Students reporting positive comments on generic proofs	9	Reduce abstraction	5
		Eliminate jargon	5
		Potential to improve student understanding	9
Students who were critical of generic proofs	9	Generic proofs are not true proofs	1
		Lack rigor	4
		Lack generality	6

Table 1. Participants' comments on generic proofs

Reduced abstraction. Five participants commented that they liked how generic proofs reduced abstraction. Students emphasized how the use of the generic example illuminated the reasoning of the proof and aided their understanding of the proof, as we illustrate below:

> G4: This [traditional proof] certainly looks more complicated because you're going to deal with arbitrary numbers like a_1, a_2, …, a_n. It's hard to make sense out of it. It's just more confusing.

Another participant noted:

> G1: It's just a lot easier to see the numbers working out and to seeing this is what we're really talking about right here. We're plugging in the numbers and this is what we really want to spit out… Rather than just talking about this general a_k's and a_n's.

Eliminated notation and jargon. Five participants also commented that generic proofs eliminated notation and jargon, emphasizing that they liked the implementation of numbers in the generic example:

> G3: I guess I could get confused in the subscripts. I definitely like it when it's like this [the generic proof], like all written down.

Further, when given the linear proof to compare the two formats, three students reported the notation and jargon used in the linear format to be intimidating and confusing. For instance, another participant said:

> G5: This [traditional proof] looks like the kinds of proofs that we had to write up that I'd always mess up with the variables or subscripts or whatever we were dealing with. I'd always lose some number or, just kind of get lost with all the different variables that we had to keep track of… This one looks scary and confusing. But this [generic proof] you can actually like, you can see what they're doing.

The potential to improve student understanding. Nine participants noted how generic proofs could improve student understanding. Seven of these students mentioned that the generic example helped them to generally understand the proof:

> G9: So it was probably easier to understand than any other proofs I've read that don't give, like an example number.

Four students also mentioned that generic proofs could be used as an aid to understanding a traditional, linear proof of the same theorem:

> G3: So it might be good just to see, oh like this [linear proof] is how to formalize it, but I think this [generic proof] is easier to understand. So maybe if you could do some combination or you can have this and that at the same time – that would be the best.

Two participants also noted that generic proofs maybe particularly helpful in proof comprehension for novice students:

> G1: But if we're focusing on education, students who are first being introduced to concepts like induction and contradiction arguments, it may be good to introduce or to transition and ease them into it using numbers as side examples… [R]ather than just posting notation all over the board and having them kind of clueless because they've never been introduced to this before.

In summary, our results are consistent with those from Rowland's (2001) survey. The students generally viewed generic proofs as a potential aid for comprehension, in part because they reduced abstraction and eliminated cumbersome notation.

Reservations about generic proofs. Nine participants were also critical of the generic proof format in some way. In general, these participants noted the possible learning gains that can be achieved from reading generic proofs, despite their criticisms. One student, G7, noted that generic proofs are not proofs, saying, "I think that it's a good first step, but I don't think it's a real proof". Four participants questioned the rigor of generic proofs. For instance, G4, "this definitely helps but it's just not a rigorous way to prove it". Both of these quotes illustrate how some participants value the potential ability of generic proofs to aid student comprehension while acknowledging they may lack mathematical rigor.

Six participants questioned whether the generic proofs were sufficiently general:

> G6: I was left with the conclusion [for the generic proof] that it works for 3. How does it work for 5? For 100? That would be my only thing, the generalizability of the proof.

These students failed to see that the reasoning in the generic proof applied to not only 3, but also all natural numbers. While each student was provided with written instruction on generic proofs, it is possible that the participants did not view the examples as generic, but as specific examples.

Quantitative Internet Study

Rationale

In the discussion of the qualitative study, we highlight that students have positive feedback after reading generic proofs based on the reduction of abstraction and the lack of notation. However, this data is limited for two reasons. First, we interviewed only ten students from the same program in the same large northeastern research university. This small sample of students reduces the generalizability of our study. Second, the finding that students had positive opinions on generic proofs does not imply that these proofs actually improve students' comprehension of them. It is possible that a proof format that is well liked by students may lead to no significant learning gains (e.g., Roy et al., 2010).

The goal of this larger quantitative internet study was to replicate the trends observed in the qualitative study and look for evidence that suggests that reading generic proofs may aid student comprehension. Specifically, we investigated students' abilities to see how a proof relates to specific examples, transfer the ideas of the proof to another theorem, summarize the proof, and see how particular statements are justified within the proof.

Methods

Participants. The participants in this quantitative internet study were recruited from mathematics majors from top universities in the United States and Canada. We recruited 106 students: students were not paid and were contacted through the secretaries of their institution's mathematics department. We provided an email to be distributed to students that explained the purpose of the experiment. We asked third and fourth year mathematics majors and minors to visit the experimental website if they would like to participate.

Materials. This study was conducted using an internet-based instrument in order to maximize our sample size. The validity and reliability of this type of study have been extensively discussed in the research methods literature (e.g. Gosling et al., 2004; Reips, 2000).

To ensure validity, we took multiple safeguards when conducting this study: 1) Each student reported whether they were seriously participating in the study, 2) the instrument recorded participant IP addresses, and 3) the instrument recorded the order in which pages of the study were viewed. Before analyzing the data, we first discarded data when there was evidence (repeated IP addresses) of a student participating in the experiment multiple times. We then removed any participant that revisited pages while completing the study or were not seriously participating. This corroborates with the methodology of Inglis and Mejia-Ramos (2009) to deal with the common threats to validity for this type of study.

The materials used for this study are included in the Appendix. For this study, we used the same generic and linear proof that was used in the qualitative study. We designed a proof comprehension assessment consisting of six items. Developed based our assessment model as discussed above (Mejia-Ramos, et al., 2012), these six questions assess students' abilities to apply the proof to examples, transfer the ideas of the proof, summarize the proof, and see how specific statements of the proof are justified within the proof.

Procedure. Each student was randomly assigned to one of two groups: 54 participants were placed in the generic group and 52 in the linear group. Participants reported their program (math major, math minor, or other) as well as their year of study (1st year undergraduate, 2nd year undergraduate, 3rd year undergraduate, 4th year undergraduate, postgraduate, or other). Next, each student was presented with instructions for the study. Participants in the generic group also received brief instruction on generic proofs, which is also included in the Appendix.

Each student was presented with a single proof; the generic or linear proof depending on their group assignment. After reading the proof participants reported, on a five-point Likert scale, how well they understood the proof and to what extent they were persuaded the claim is true given the information presented in the proof. Students were also asked whether they found the result applied generally to any natural number n and whether they believed the proof was valid. The participants were next presented with the six assessment questions in a randomized order. Finally, students were asked to report on a five-point Likert scale whether they liked the format in which the proof was presented and given space to give any additional comments.
Each question appeared on a new screen and participants were asked not to move back in their browser to review the proof or change their answers for previous questions. Participants were informed that if they did such their data would not be considered for analysis and were warned if they clicked to progress to the next page without answering all of the questions. Subsequently, we removed any participant from our analysis if they disobeyed these guidelines.

Results

Comprehension assessment results. The participants' performance on the comprehension test is presented in Table 2. We note two trends in the data. First, the generic group performed better on the two example questions. Second, the generic group performed worse on three of the four questions that did not pertain to examples. That the generic group performed worse on the transfer question is the opposite effect of what Malek and Movshovitz-Hadar (2011) found in their study.

Question	Generic Group (N=54)	Linear Group (N=52)
Example 1	91%	83%
Example 2	94%	85%
Transfer	30%	43%
Summary	65%	62%
Justification 1	37%	63%
Justification 2	74%	87%

Table 2. Participants' performance on the comprehension test.

In Figure 1, we categorize the assessment items by example items and non-example items. In this analysis, the generic group performed statistically reliably worse than the linear group on non-example item with $t(104)=2.33$ and $p<.05$ ($p=.0216$). The data also suggests that the generic group outperformed the linear group on example items, although not statistically reliably so.

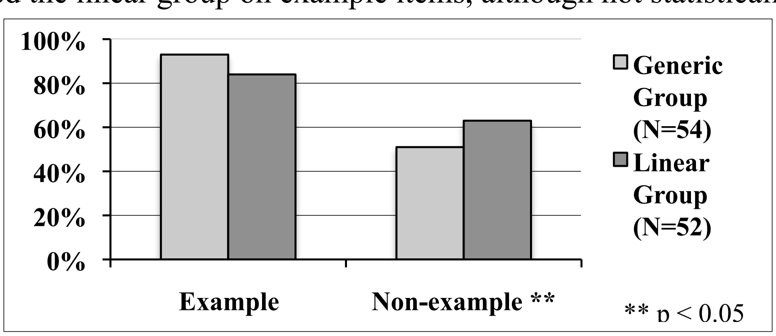


Figure 1. Participant performance on example and non-example items.

Participant opinions on generic proof format. Of the 54 participants in the generic group, 38 claimed to like the format, 12 disliked the format, 3 were neutral, and 1 did not respond. Of the 38 who reported favorably towards the generic proofs, 20 left comments. Eight of those participants commented on why they liked the format reporting on the reduced abstraction and use of examples. It is interesting to note that there were 12 participants, who while having a favorable view of the generic proof format, still had reservations about the validity of a generic proof or were left wanting a more rigorous proof, as we illustrate below:

> Although I really like the idea of illustrating the formal proof with specific examples, it is no substitute for a formal proof by induction. Examples are excellent tools and should be used when writing/reading proofs, but specific examples do not prove a theorem. A formal inductive proof is needed.

This suggests that some participants may appreciate the value of generic proofs in respect to comprehension, yet still value deductive proofs for other purposes, such as validity and generality.

Of the 12 participants who responded unfavorably to the generic format, ten participants responded that they preferred a rigorous proof, were not convinced by the proof, or found the generic example to be unnecessary.

These results are again consistent with Rowland (2001) as well as our qualitative study; the illustrate that students generally have a positive opinion of generic proofs, although even some students who view these proofs favorably value some features of a deductive proof that generic proof lacks.

Generality and validity. When analyzing the students' reports on their beliefs of generality and validity, we found some noteworthy results. Both the linear group and the generic group had the same views about how general their arguments were, with 17% of the linear group and 18% of the generic group questioning whether the proofs were sufficiently general. However, the generic group (40%) was significantly more likely than the linear group (17%) to challenge the validity of the proof (Fisher exact, p=.018).

Discussion

Summary of Findings

In the qualitative study, the participants indicated that they appreciated that generic proofs reduced abstraction and eliminated cumbersome notation, thereby having the potential to improve comprehension. The quantitative study corroborated these findings. In the quantitative study, the use of a generic proof significantly hindered student comprehension with respect to assessment items that did not involve applying the ideas of the proof to specific examples.

The results of this study suggest that although students may believe that generic proofs improve comprehension, they do not actually help and possibly hinder their abilities to answer assessment questions relating to justification and transfer. To avoid misinterpretation, we are *not* suggesting that generic proofs are not a useful pedagogical tool for improving proof comprehension. We believe it is quite possible that if generic proofs were introduced to students in a different way that other learning gains might be observed. What this study does suggest is that simply giving students generic proofs may not improve performance and possibly may hinder performance. However, as this study was conducted with only one proof, we would need to replicate this study with more proofs to gain confidence in the generality of our findings.

Inaccurate Student Beliefs about Their Learning from Generic Proofs

Our data suggest that although students may prefer or value generic proofs, they may not be entirely beneficial to students. The generic group generally had positive opinions on the generic proofs and reported that generic proofs could aid student comprehension. Yet average student performance on a non-example assessment items was significantly worse than that of the linear group. One way to account for this data is that students preferred generic proofs since it eliminated the requirement to cope with abstraction and technical jargon; however, although grappling with the abstraction and notation of traditional proofs might be difficult and unpleasant, a lot may be learned from engaging in this process. Another account is based on Selden and Selden's (2003) observation that mathematics majors tend to focus on calculations when reading a proof while ignoring its global structure. If so, generic proofs would be preferable to students as they made these calculations easier, but may actually wind up accidentally encouraging the unproductive behavior that Selden and Selden highlighted.

Significance of Results

Our findings contribute to the current literature in three ways: Our data corroborate Rowland's (2001) claim that students favor generic proofs, challenge Malek and Movshovitz-Hadar's (2011) claim generic proofs improving student comprehension, and contribute to the growing literature on empirical studies on the efficacy of alternative proof formats.

As noted in the previously, Rowland (2001) suggested that students had a favorable impression of generic proofs. In both our qualitative and quantitative studies, we also found that students generally had a favorable opinion of generic proofs, in part because they reduced abstraction, eliminated jargon, and improved understanding. Nonetheless, these students still noted that generic proofs may lack the rigor and generality that a traditional proof provides.

Malek and Movshovitz-Hadar (2011) claimed that generic proofs significantly improve student comprehension, specifically in the area of transfer. However, we found in our quantitative study that students who read the generic proof performed worse on transfer questions. More generally, the linear group outperformed the generic group on non-example assessment questions. More research is needed to see whether the discrepancy between our results is due to the small sample size used in Malek and Movshovitz-Hadar's (2011) study, the different mathematical domains being investigated (linear algebra vs. number theory), or something else.

Finally, we noted in the introduction that there are few empirical studies on the alternative proof formats suggested in the literature. We contribute this study on generic proofs to this collection. Most importantly, like the two previous quantitative studies of proof comprehension (Fuller et al., 2011; Roy et al., 2010), we failed to find evidence that reading an alternative version of a proof led students to understand the proof better than reading a traditional version of the same proof. There are two possible accounts to this data. The first is that there is no panacea for curing students' difficulties with proof comprehension and simply presenting proofs in a different format will not lead to large learning gains. The second is that alternative proof formats can lead to learning gains, but students need more preparation prior to studying these proofs or more experience with these formats for these gains to be realized. An important possibility that we did not consider is that generic proofs would improve comprehension if presented prior to a formal proof, rather instead of a formal proof. We note that although researchers suggest alternative formats for proof presentation, they often do not describe, at a prescriptive level, how these formats should be used. Future research that investigates specific classroom situations that meaningfully incorporate these alternative proof formats and systematically documents learning gains would therefore not only serve to demonstrate the potential of these formats, but also provide insight for what it takes to use these formats effectively.

References

Alcock, L. (2009). e-Proofs: Students experience of online resources to aid understanding of mathematical proofs. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education.*

Anderson, J., Boyle, C., & Yost, G. (1986). The geometry tutor. *Journal of Mathematical Behavior*, *5*, 5-20.

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

Fuller, E., Mejia-Ramos, J.P., Weber, K., Samkoff, A. Rhoads, K., Doongaji, D., & Lew, K. (2011). Comprehending Leron's structured proofs. In *Proceedings of the 14th Conference for Research in Mathematics Education.*

Gosling, S.D., Vazire, S., Srivastava, S., & John, O.P. (2004). Should we trust web-based studies? A comparative analysis of six studies about internet questionnaires. *American Psychologist*, *59*, 93-104.

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

Harel, G. (1998). Two dual assertions: The first on learning and the second on teaching (or vice-versa). *American Mathematical Monthly, 105,* 497-507.

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

Inglis, M. & Mejia-Ramos, J.P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, *27*, 25-50.

Kline, M. (1973). *Why Johnny can't add: The failure of New Mathematics*. New York, NY: St. Martin's Press.

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly, 90(3),*174-184.

Malek, A. & Movshovitz-Hadar, N. (2011). The effect of using Transparent Pseudo-Proofs in linear algebra. *Research in Mathematics Education, 13,* 33-58.

Mantini, L., Mills, M., & Johnson, J. (2012). Reading comprehension of series convergence proofs in Calculus II. Paper presented at the 15th Conference for Research in Undergraduate Mathematics Education. Portland, OR.

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.

Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. In Proceedings of the *14th Conference for Research in Undergraduate Mathematics Education.*

Porteous, K. (1986). Children's appreciation of the significance of proof. *Proceedings of the Tenth International Conference of the Psychology of Mathematics Education* (pp. 392-397). London, England.

Reips, U.-D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 89–117). San Diego: Academic Press.

Rowland, T. (2001). Generic proofs in number theory. In S. Campbell and R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction.* (pp. 157- 184). Westport, CT: Ablex Publishing.

Roy, S., Alcock, L., & Inglis, M. (2010). Supporting proof comprehension: A comparison study of three forms of comprehension. In *Proceedings of the 13th Conference for Research in*

Undergraduate Mathematics Education.
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.
Strauss, A. & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* London: SAGE.
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 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. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning,* 12, 306-336.

Appendix: Instructions, generic proof, linear proof, and questions (final version from quantitative study)

Instructions on generic proof

In college math classes, theorems are traditionally stated and then proven in general, abstract terms. Some mathematicians have suggested another way of presenting proofs for theorems — by illustrating the proof with one or more specific examples. For instance, to justify a fact about the first n odd natural numbers, one might illustrate how the proof works with the first 3 odd natural numbers. Though the reasoning might only be shown for a certain set of examples, it would work in a similar way for any set of examples. One goal of this study is to see how students read proofs presented in this manner.

Generic Proof

Theorem. Let S_n be the set of ordered finite sequences of positive integers that sum to n. (For example, $(2,1,2),(1,2,2)$, and (5) are distinct elements of S_5). Then the number of elements in S_n is 2^{n-1}. In other words, there are 2^{n-1} ways to express n as an ordered sum of positive integers.

Proof: (By induction).
Base case. If $n = 1$, the only element in S_1 is (1). Hence there are $1 = 2^0 = 2^{1-1}$ elements of S_1.
Inductive case. Assume that the theorem is true for $n = k$; that is, S_k has 2^{k-1} elements. We will illustrate how each element of S_k generates two new elements of S_{k+1} using the example $k = 3$.

> For each element of S_3, we generate an element of S_4 either by increasing the last entry in the sequence by 1 or by appending a 1 at the end of the sequence.
>
> To illustrate, $(2,1)$ is an element of S_3 and generates two elements of S_4. We add 1 to the last entry to get $(2,2)$ or we append a 1 at the end of the sequence to get $(2,1,1)$. Using this process,
> (3) generates (4) and $(3,1)$
> $(1,2)$ generates $(1,3)$ and $(1,2,1)$
> $(2,1)$ generates $(2,2)$ and $(2,1,1)$
> $(1,1,1)$ generates $(1,1,2)$ and $(1,1,1,1)$
> Each generated element is in S_4.
>
> Next, we show each element of S_4 was generated from one element of S_3. If the last entry of the sequence in S_4 is a 1, simply eliminate the 1 from the sequence. For instance, $(1,2,1)$ is in S_4 and was generated by $(1,2)$ in S_3.
>
> Otherwise, if the last entry of the sequence in S_4 is greater than 1, decrease the last entry of the sequence by 1. For instance, $(2,2)$ is in S_4 and was generated by $(2,1)$ in S_3.
>
> Since every element in S_3 generates two elements in S_4 and every element of S_4 is generated by exactly one element of S_3, there are twice as many elements in S_4 as there are in S_3.

The logic illustrated above can be applied to any S_k and S_{k+1}. Thus, S_{k+1} always has twice as many elements as S_k. By the inductive hypothesis, there are 2^{k-1} elements in S_k. Therefore, there are 2^k elements in S_{k+1}.

Linear Proof

Theorem. Let S_n be the set of ordered finite sequences of positive integers that sum to n. (For example, $(2,1,2),(1,2,2)$, and (5) are distinct elements of S_5). Then the number of elements in S_n is 2^{n-1}. In other words, there are 2^{n-1} ways to express n as an ordered sum of positive integers.

Proof: (By induction).
Base case. If $n = 1$, the only element in S_1 is (1). Hence there are $1 = 2^0 = 2^{1-1}$ elements of S_1.
Inductive case. Assume that the theorem is true for $n = k$; that is, S_k has 2^{k-1} elements. We show that each element of S_k generates two elements of S_{k+1}.

Let $(a_1, a_2, \ldots, a_m)$ be an element of S_k. By definition of S_k, we know that $a_1 + a_2 + \cdots + a_m = k$.

We generate two elements of S_{k+1} as follows. The first one is generated by increasing the last term of the sequence by 1 to yield $(a_1, a_2, \ldots, a_m + 1)$. This is an element of S_{k+1} since

$$a_1 + a_2 + \cdots + (a_m + 1) = (a_1 + a_2 + \cdots + a_m) + 1 = k + 1.$$

The other element is generated by appending a 1 at the end of the sequence to yield $(a_1, a_2, \ldots, a_m, 1)$. This is also an element of S_{k+1}.

Next, we show that each element of S_{k+1} was generated from one element of S_k. Let $(b_1, b_2, \ldots, b_m)$ be an element of S_{k+1}.

If $b_m = 1$, simply eliminate b_m from the sequence. In this case, $(b_1, b_2, \ldots, b_m)$ was generated by $(b_1, b_2, \ldots, b_{m-1}) \in S_k$.

Otherwise, if $b_m > 1$, decrease the last entry of the sequence by 1. In this case, $(b_1, b_2, \ldots, b_m)$ was generated by $(b_1, b_2, \ldots, b_m - 1) \in S_k$ [note that $b_m - 1$ is a positive integer because $b_m > 1$].

Since every element of S_{k+1} is generated by exactly one element of S_k, this ensures that the process above yields S_{k+1} and does not double count.

Because every element of S_k generates two elements in S_{k+1}, S_{k+1} has twice as many elements as S_k. By the inductive hypothesis, there are 2^{k-1} elements in S_k. Thus, there are 2^k elements in S_{k+1}.

Questions

Summary

Which of the options below best expresses the main ideas used in the proof?

A. Each element of S_k generates two elements of S_{k+1}, because we can add 1 to the last entry in the sequence or append a 1 at the end of the sequence. This generates all elements of S_{k+1}.

B. *Base case.* If $n = 1$, the only element in S_1 is (1). Hence there are $1 = 2^0 = 2^{1-1}$ elements of S_1. *Inductive case.* Suppose S_k has 2^{k-1} elements. Then S_{k+1} has 2^k elements.

C. S_{k+1} is obtained from sequences of S_k either by removing the last entry in the sequence or by subtracting a 1 from the last entry. Hence, every element of S_{k+1} generates two elements of S_k.

D. I don't know.

Justification 1

How do we know that the method used in this proof did not produce the same element of S_{k+1} twice?

A. Each element of S_k produces exactly two elements of S_{k+1}.

B. We can tell which unique element of S_k produced a particular element of S_{k+1}.

C. Adding 1 to the last entry of a sequence cannot yield the same result as appending a 1 at the end of the sequence.

D. I don't know.

Justification 2

If (a, b, c, d) is an element of S_k, how do we know that $(a, b, c, d, 1)$ is an element of S_{k+1}?

Example 1

$(2, 3, 1)$ is an element of S_6. Using the ideas of the proof, what two elements of S_7 can be formed from $(2, 3, 1)$?

Example 2

$(3, 1, 2)$ is an element of S_6. Using the ideas of the proof, what element of S_5 was used to form $(3, 1, 2)$?

Transfer

Consider the following theorem: There are 2^n subsets of $\{1, 2, \dots, n\}$.

Two students came up with the following approaches to prove this theorem. Which of their approaches is most consistent with the ideas of the proof?

A. Take a subset $\{b_1, b_2, \dots, b_m\}$ of $\{1, 2, \dots, k\}$. This forms two subsets of $\{1, 2, \dots, k+1\}$ as follows. Take the original subset $\{b_1, b_2, \dots, b_m\}$ or include $k+1$ in this subset to yield $\{b_1, b_2, \dots, b_m, k+1\}$.

B. Take a subset $\{b_1, b_2, \dots, b_m\}$ of $\{1, 2, \dots, k\}$. This forms two subsets of $\{1, 2, \dots, k+1\}$ as follows. Add 1 to the last element of $\{b_1, b_2, \dots, b_m\}$ to obtain $\{b_1, b_2, \dots, b_m + 1\}$ or append a 1 at the end of the subset to obtain $\{b_1, b_2, \dots, b_m, 1\}$.

C. I don't know.

HOW AND WHY MATHEMATICIANS READ PROOFS: AN EXPLORATORY QUALITATIVE STUDY AND A CONFIRMATORY QUANTITATIVE STUDY

Keith Weber Juan Pablo Mejia-Ramos
Rutgers University

Abstract. In this paper, we investigate how and why mathematicians read the published proofs of their colleagues. Based on qualitative interviews with nine mathematicians, we posit that mathematicians understand proofs in three ways: as cultural artifacts with a social history, as a sequence of inferences, and as the application of methods. Each type of understanding is based, at least in part, on non-deductive evidence. A survey with 118 mathematicians confirms the generality of these findings. We conclude by arguing that more comprehensive frameworks for how mathematicians gain conviction are needed.

Key words: Mathematical Practice; Proof comprehension; Proof reading

In the last thirty years, there has been a great deal of research in mathematics education in the area of justification and proof. Although this research has focused on a wide range of topics, including the construction of proof, the epistemological nature of proof, and the development of proof in mathematical classrooms, only recently has research in mathematics education investigated the *reading* of mathematical proof (e.g., Selden & Selden, 2003).

Most researchers who have examined the reading of proofs have sought to understand students' conceptions of proofs by asking them to read different types of arguments and to *evaluate* these arguments against a given set of criteria (e.g. personal preference, persuasiveness, mathematical validity). This literature has produced three main findings. First, many students find empirical arguments to be personally convincing and representative of the ways that they would justify mathematical statements (e.g., Martin & Harel, 1989; Healy & Hoyles, 2000). Second, mathematics majors often have difficulty distinguishing between valid mathematical proofs and flawed mathematical arguments (e.g., Selden & Selden, 2003; Alcock & Weber, 2005). Third, students may view a deductive proof of an assertion merely as evidence in favor of the assertion rather than necessitating its truth (e.g., Fischbein, 1982). Together, these findings suggest that students become convinced of mathematical assertions for different reasons than mathematicians do (e.g., Harel & Sowder, 1998).

A goal of many research programs is to lead students to think and behave more like mathematicians with respect to proof. Harel and Sowder (2007) explicitly contend that the purpose of mathematics instruction should be "to help students gradually develop an understanding of proof that is consistent with that shared and practiced in contemporary mathematics". To accomplish this goal, some researchers have conducted teaching experiments designed to have students develop the same standards of conviction as mathematicians (e.g. Harel, 2001) while others have created learning environments designed to have students engage in proof-related activity that is similar to mathematicians' practice (e.g., Lampert, 1990; Maher, Muter, & Kiczek, 2007).

If goals of mathematics instruction include having students (1) engage in the same types of proof-related activities that mathematicians do, (2) behave like mathematicians in these activities, and (3) adopt mathematicians' beliefs regarding proof, then it is necessary to have an accurate understanding of the types of activities that mathematicians engage in, how mathematicians perform them, and what their beliefs about proof actually are. The RAND

Mathematics Study Panel (2003) concluded that more research on mathematicians' practice pertaining to justification and proof is needed to form a sufficient basis to design instruction. Consistent with this report, several recent empirical studies and philosophical theses have investigated this topic and yielded surprising findings about mathematicians' practice with justification and proof (e.g., Rav, 1999; Inglis, Mejia-Ramos, & Simpson, 2007; Weber, 2008; Inglis & Mejia-Ramos, 2009) and, in some cases, these findings have had important implications for the teaching of mathematics (e.g., Hanna & Barbreau, 2008). With regards to reading proofs, Konoir (1993) contended that "getting to know the complex processes and mechanisms of reading text has essential significance for the didactics of mathematics" (p. 251), noting that few such studies have been conducted and arguing that studies of mathematicians' practice when reading proof need to be conducted.

The goal of this paper is to contribute to the mathematics education community's understanding of mathematicians' practice with regard to the reading of mathematical proof by addressing the following questions:

- For what purposes do mathematicians read the proofs of their colleagues?
- How do they read proofs to achieve these aims?
- What role does non-deductive reasoning play in this proof reading?

Theoretical Perspective

Our investigation is informed by two theoretical frameworks. The first is an extension of the warrant typology of Inglis, Mejia-Ramos, and Simpson (2007). Inglis et al. investigated the different sources of information, or warrants, that mathematicians may use to gain conviction that a mathematical statement is true. When an assertion claims every element of a set satisfies a given property, one may check that this assertion is true by verifying that a proper subset of the given set satisfies this property. We refer to this type of argument as *empirical*. One may increase one's confidence that a claim is correct because an authoritative source endorsed that claim. We refer to this type of argument as *authoritative*. One may increase one's conviction that a statement is true because it is consistent with one's representations or mental models of the involved concepts. We refer to this type of argument as *structural-intuitive*. Finally, one may gain confidence in an assertion by producing or observing a deductive argument that derives a particular assertion. We refer to this type of justification as *deductive*.

The second theoretical framework that we adopt is an extension of Boero's (1999) stages of proving. Boero argues that proof production by mathematicians usually has five stages: (i) generating a conjecture, (ii) formally stating the conjecture, (iii) exploring the conjecture, (iv) developing an informal justification, and (v) logically chaining the informal justification into a proof. Boero (1999) noted that although empirical warrants (and other informal reasoning) do not appear in the products of proof production (namely (ii) and (v)), they play a significant role in the generation of these products (stages (i), (iii), and (iv)). In other words, one cannot "prove by example" or "prove by picture", but examples and pictures are critical in the formation of conjectures and proofs.

We contend that the mathematical work of generating a proof does not end in stage (v). After producing a proof, (vi) the mathematician submits this work for review by other mathematicians, (vii) other mathematicians evaluate the argument, and if the review is positive, (viii) other mathematicians read and learn from the argument in a published source. Our paper primarily concerns stage (viii)—that is, how do mathematicians evaluate and comprehend the published proofs of others—and what warrant types mathematicians use to make these evaluations.

Qualitative study

In addition to the writings of mathematicians and philosophers of mathematics, we will use two primary sources of data to generate our model of how and why mathematicians read proofs. The first comes from a paper (Weber, 2008) in which the first author of this paper (a) analyzed the way in which eight research active mathematicians determined if mathematical arguments were correct and (b) interviewed these eight mathematicians about their professional practice in evaluating the proofs of others. A central result from this paper was that mathematicians occasionally marshaled non-deductive evidence to bridge what they perceived to be inferential gaps in the proof. In particular, these mathematicians often checked if particular statements in a proof were true by verifying that they held for carefully chosen examples. The second source of data, presented in this paper, comes from interviews with nine other mathematicians on their professional practice of reading the published proofs of others.

As the data from this study comes from a total of 17 mathematicians, we recognize its generalizability is limited. To compensate for this shortcoming, we will also present a quantitative study in which we survey 118 mathematicians on whether they agree with the hypotheses that we generated from our qualitative studies.

Research methods

Nine professional mathematicians participated in this study and agreed to meet individually with the first author for a semi-structured interview. All participants were tenured mathematics professors at a large research university in the northeast United States. All were highly successful researchers in their fields of study, which included analysis, algebra, and differential equations. Each interview was semi-structured and was one to two hours long. Each interview was audiotaped and then transcribed. The goal of the interview was to investigate the reasons why and the ways in which the mathematicians read the published proofs of their colleagues. The analysis in this paper will focus on the participants' responses to the following questions:

- In your own mathematical work, I assume you sometimes read the published proofs of others. What do you hope to gain by reading these proofs?
- What do you think it means to understand a proof?
- What are some of the things that you do to understand proofs better?
- Does considering specific examples ever increase your confidence that a proof is correct?

The data were analyzed using an open coding scheme (Strauss & Corbin, 1990). To parse the data, we first defined an episode as an instance when a participant either discussed (a) a reason for reading proofs, (b) a strategy for reading proofs, or (c) a justification for either (a) or (b). An initial description was given to each episode. Similar episodes were given preliminary category names and definitions. New episodes were placed into existing categories when appropriate, but also used to create new categories or modify the names or definitions of existing categories. This process continued until a set of categories was formed that was grounded to fit the data.

Results

All mathematicians confirmed that reading the published proofs of others was a significant part of their mathematical practice. In fact, M7's initial response to whether he read other mathematicians' proofs was: "What do you think I was doing when you came into my office?"

Do mathematicians check proofs for correctness?

Six participants indicated that when reading proofs in mathematical journals, they would sometimes do so to determine if the proofs were correct (although some of these participants may have been referring to refereeing a paper, rather than reading a published paper). One representative response was included below:

I: What do you hope to gain when you read these proofs?

M4: Okay. Two things. One is I would like to find out whether their asserted result is true, or whether I should believe that it's true. And that might help me, if it's something I'd like to use, then knowing it's true frees me to use it. If I don't follow their proof then I would be psychologically disabled from using it. Even if somebody I respect immensely believes that it's true. More importantly, I want to understand the proof technique in case I can use bits and pieces of that proof technique to prove something that they haven't yet, that the original author hasn't yet proved.

Two of the participants specifically claimed that they did *not* read proofs to check their correctness. M8 emphasizes this point in the excerpt below.

M8: Now notice what I did *not* say. I do not try and determine if a proof is correct. If it's in a journal, I assume it is. I'm much more interested in the ideas of the proof.

Similarly, when asked what he hoped to gain out of a proof, M6 did not specifically mention determining if a proof was right, prompting the interviewer to ask why this was not said.

I: One of the things you didn't say was you would read it to be sure the theorem was true. Is that because it was too obvious to say or is that not why you would read the proof?

M6: Well, I mean, it depends. If it's something in the published literature then… I've certainly encountered mistakes in the published literature, but it's not high in my mind. So in other words I am open to the possibility that there's a mistake in the proof, but I… it's not… [pause]

I: But you act on the assumption that it's probably correct?

M6: Yeah, that's right. That's right.

M6 and M8 both act on the assumption that proofs published in journals are probably correct and do not feel the need to personally validate them. Among the nine mathematicians who were interviewed, these two mathematicians were the only ones to make a comment of this type, although other participants may have held similar viewpoints. Several mathematicians interviewed by Weber (2008) also indicated that they would not check a proof for correctness because they trusted the proof if it appeared in a reputable journal.

Reading a proof for insight.

All nine mathematicians claimed they read proofs for ideas that might be useful in their own research. One instance is provided above, in which M4 says that this is more important than checking for correctness. Other excerpts are provided below:

M1: Theorems as a way of organizing major results are extremely useful, but they involve a decision on the part of the person who's writing the statement of the theorem of what thing to take as a hypothesis, what is the conclusion, and what route to follow to get from hypothesis to conclusion. And often along that route there are techniques that could have been stated as separate theorems but are not, and then you read the proof carefully and *you discover these are things that you can use*. That's certainly from a pragmatic point of view, *that's an important part of reading proofs, that you steal good ideas out of good proofs.*

M6: Well, I would say most often is to get some ideas that might be useful to me for proving things myself. [When asked to elaborate] Ok, actually, I'll say two things. One is, it's just to satisfy my innate curiosity as to *what are they doing to get this conclusion*. So that's one aspect. But then the other aspect is… I mean, usually I'm reading something because it seems to have some connection to some problems that I was interested in. *I'm hoping that if that the tools they're using or ideas they're using might connect up to some of these problems that I have thought about.* (In both excerpts, italics were our emphasis).

The goal of these mathematicians' reading seems to go beyond comprehending the proof. They seek to find ideas and techniques in their proofs that will help them address their own research questions. In the following excerpt, I asked M5 to go into more depth about what he meant with respect to looking at the ideas of a proof and seeing if they could apply elsewhere.

M5: As a researcher, I want to understand the idea of the proof and to see if that idea could be applied elsewhere.

I: The second point that you made, the one about ideas, is something that I've been hearing from your colleagues too. Can you elaborate on that?

M5: Sure. Sometimes when a mathematician answers a hard question, he has a new way of looking at the problem or a new way of thinking about it. As a researcher, when you see this, sometimes you can use this idea to solve problems that you are working on. Let me give you an example. We were having trouble showing bounds for approximation techniques on this space with an unusual norm. Someone realized that you could use this particular partial differential equation to find these bounds. This new idea made a lot of the other problems easier. The idea wasn't easy. It wasn't obvious at all that this partial differential equation was relevant. That was a great insight. But once we had the idea, it allowed us to approach questions that were inaccessible before.

I: So after this theorem came out, a lot of other theorems were proved using this idea?

M5: Oh yeah. But it doesn't always have to be big things, although this one was. Sometimes when I read a proof, I get an idea that helps me get around a little thing that I was stuck with.

Not only did the participants consistently emphasize that ideas were the primary reason that they read proofs, but two participants went so far as to question whether a proof of a new theorem is of value if it does not contain new ideas.

M9: [As editor of a journal,] I occasionally get papers where the author took an idea from a new proof that just came out, took the idea from the proof, and applied it in a straightforward fashion to prove some new theorem. I'm reluctant to publish these types of papers.

M1: When I was on the editorial board for one of the journals, one of the instructions we had was "it's not allowed to just publish a paper where you've taken somebody else's proof and simply made a different statement of what we should get out of the proof". That if there is a really good theorem and you come up with an original, alternate proof of that theorem, that could be publishable. But just to take the same proof and say well, we can state the conclusion differently. That's not considered professionally acceptable as a result … Mathematics has this very high, perhaps unrealistically high standard for what is admissible as a claim of research.

These excerpts (which speak to reviewing a proof rather than reading a published proof) suggest that the result of a mathematical piece of research does not depend solely, perhaps not even primarily, on the significance of the theorem being proved and the validity of its proof. The originality and utility of the ideas in the proof is of crucial importance. This notion is endorsed by Rav (1999), who emphasizes that proofs, not theorems, are the bearers of mathematical knowledge (p. 20). Like M1, Rav suggests that what a mathematician claims a proof establishes is somewhat arbitrary; it is the proof method that is of primary importance: "Think of proof as a network of roads in a public transportation system, and regard statements of theorems as bus stops; the site of the stops is just a matter of convenience" (p. 20-21).

Understanding a proof.

When asked what it means to understand a proof, all nine participants indicated that understanding did not solely consist of knowing how each step followed logically from previous steps. In fact, several participants distinguished between understanding a proof logically and understanding the central ideas of a proof.

> M8: There are different levels of understanding. One level of understanding is knowing the logic, knowing why the proof is true. A different level of understanding is seeing the big idea in the proof. When I read a proof, I sometimes think, how is the author really trying to go about this, what specific things is he trying to do, and how does he go about doing them. Understanding that, I think, is different than understanding how each sort of logical piece fits together.

In fact, M5 explained that although a full understanding of a proof involved verifying that each particular instance within a proof was valid, this was a process that he often did not engage in.

> M5: [To understand a proof] means to understand how each step followed from the previous one. I don't always do this, even when I referee. I simply don't always have time to look over all the details of every proof in every paper that I read. When I read the theorem, I think, is this theorem likely to be true and what does the author need to show to prove it's true. And then I find the big idea of the proof and see if it will work. If the big idea works, if the key idea makes sense, probably the rest of the details of the proof are going to work too.

The use of examples in understanding a proof.

When asked what they did to understand a proof better, six participants claimed that they would consider how the proof related to specific examples.

> M4: Commonly, if I'm really befuddled and if it's appropriate, I will keep a two-column set of notes: one in which I'm trying to understand the proof, and the other in which I'm trying to apply that technique to proving a special case of the general theorem.

> M1: I'm doing a reading course with a student on wallpaper groups and there is a very elegant, short proof on the classification of wallpaper groups written by an English mathematician. So in reading this… so this is one where he's deliberately not drawing pictures because he wants the reader to draw pictures. And so I'm constantly writing in the margin, and trying to get the student to adopt the same pattern. Each assertion in the proof basically requires writing in the margin, or doing an extra verification, especially when an assertion is made that is not so obviously a direct consequence of a previous assertion.
>
> I: So you're writing a lot of sub-proofs?
>
> M1: I write lots of sub-proofs. And also I try to check examples, especially if it's a field I'm not that familiar with, I try to check it against examples that I might know.

When asked if they ever used examples to increase their confidence that a proof is correct, all nine participants emphatically answered yes. Indeed, to some participants, this question was almost meaningless since they claimed that they never read a proof without considering examples. Many of the participants discussed using examples so they could view the proof as a generic proof, as M4 does in the excerpt above. When asked if he used examples to increase his confidence in the correctness of a proof, M5 responded:

> M5: Always. Always. Like I said, I never just read a proof at an abstract level. *I always use examples to make sure the theorem makes sense and the proof works.* I'm sure there are some mathematicians that can work at an abstract level and never consider examples, but I'm not one of them. *When I'm looking through a proof, I can go off-track or believe some things that are not true.* I always use examples to see that makes sense. (Italics were our emphasis)

The italicized portions of the excerpt above illustrate how some mathematicians claim not to be able to work on an abstract, or purely logico-deductive, level. Due to human error, even professional mathematicians can "go off track and believe some things that are not true". As Thurston (1994) notes, mathematicians "are not good at checking *formal correctness* of proofs, but they are quite good at detecting potential weaknesses or flaws in proofs" (p. 169, emphasis is the author's). Checking the logic with examples and other forms of background knowledge (M1 mentions the construction of diagrams above) appears crucial for some mathematicians to reliably understand and validate a proof.

A model for how mathematicians read proofs

Based upon the interview data and the philosophical literature, we posit that mathematicians may understand a proof in three different ways: as *a cultural artifact*, a *sequence of inferences*, or as *the application of methods* (Rav, 1999). The way in which a proof is viewed influences how the proof is read and evaluated.

Proof as cultural artifact.

In order to evaluate the validity of a given argument, at least some mathematicians rely on evidence that is not directly related to the content of the argument, but rather the contextual history the argument underwent to be sanctioned as a proof by the mathematical community. For instance, in this paper, we present two mathematicians who claimed that when they read a proof in a journal, they act on the assumption that the proof is correct. Similarly, in Weber (2008), one mathematician claimed, "to be honest, when I read papers, I don't read the proofs … if I'm convinced that the result is true, I don't necessarily need to read it, I can just believe it". In these cases, the mathematicians appear to be saying that since *other mathematicians* checked the proof and claimed it was valid, they were prepared to believe the proof was valid.

This does not appear to be atypical. For instance, Jackson (2006) described one mathematician who believed Perelman's proof of the Poincare conjecture "must be right" because (i) if it was not, the collective expertise of the mathematical community would have found the mistake and (ii) Perelman's work had been reliable in the past (p. 899). His evaluation of the validity of Perelman's proof did not come from the deductive process that Perelman employed, but from the authority of the mathematical community and Perelman himself (cf., Inglis & Mejia-Ramos, 2009). Note that we are not claiming that mathematicians will believe a mathematical assertion is true because an expert in their field *claimed* it was so; rather they will believe a proof is valid because it was *validated* by mathematicians who presumably have the expertise to locate a fault in the proof if one existed.

Proof as a sequence of inferences.

As one would expect, in order to *comprehend* or *evaluate the validity* of a given argument, mathematicians seem to use what is commonly described as line-by-line reading (see Weber, 2008) where one verifies that each non-trivial assertion in a proof is a logical consequence of previous assertions.

When evaluating the validity of a given proof, a mathematician may (intuitively and implicitly) assign a probability p_i to his or her level of confidence that the i^{th} inference of the proof is correct. The probability that every inference in a proof is correct—i.e., that the proof is fully valid—is then the product of each of the probabilities assigned to each inference, that is ($p_1p_2p_3\ldots p_n$). If the proof is short and in a relatively simple domain, it is possible to become (nearly) absolutely certain that every step within a proof is correct. However, as DeMilo, Liptus, and Perlis (1979) argue, for lengthy proofs in complex domains, there will be a non-trivial probability that some of the assertions in the proof do not follow validly from previous claims. Indeed, Davis (1972) and Hanna (1991) claim that half of the published proofs contain logical errors and many proofs are rife with errors.

One way to increase one's confidence that a proof is correct is to find inferences in the proof that are problematic (i.e., have a probability value below a certain threshold) and examine them more closely to increase one's confidence in them. For instance, suppose that one only assigned a value of 0.9 that the third assertion of a proof was correct. A mathematician might construct a sub-proof with the third assertion as claim (as described above), where his or her confidence in each of the sub-proof's inferences is very high. This would be using deductive evidence to increase one's overall probability that the original proof does not contain any logical flaws. However, this is not always how mathematicians behave. Weber (2008) observed that mathematicians occasionally used empirical evidence to increase their confidence about a particular inference within a proof. For instance, one participant was uncertain about a particular assertion within a proof, but gained enough confidence to judge that assertion (and the proof in its entirety) as valid after verifying the assertion held for a single example. There were other instances in Weber's study where mathematicians increased their conviction of particular assertions within a proof because they noticed a pattern, searched for a counterexample but failed to find one, or produced a generic proof. These results are corroborated by the data reported in this paper. Although no mathematician claimed to verify an assertion only by looking at an example, all claimed that inspecting examples increased their confidence that a proof is correct. As M5 noted, "When I'm looking through a proof, I can go off-track or believe some things that are not true. I always use examples to see that makes sense."

Examples seemed particularly useful for increasing one's understanding of the proof. Six of the nine mathematicians in this study claimed to use examples to help them understand a proof. Indeed, several of them described a process by which the line-by-line reading of the proof would be accompanied by a parallel study of one or more specific examples.

Proof as the application of methods.

We do not believe that mathematicians' comprehension of a proof amounts to their comprehension of each step in the proof, or that their confidence in a theorem is equivalent to their confidence that every step within its proof is valid. Extracting ideas that could be useful in their own research (a process that mathematicians in this study considered to be crucial) seems to be more sophisticated than what we have described as line-by-line reading. Similarly, it seems unlikely that proofs are normally checked solely by appeals to authority and line-by-line reading. Davis (1972) and Hanna (1991) estimate that half of the published proofs in mathematics contain logical errors, but also argue that most of the theorems published in the literature are true. This would imply for an arbitrary theorem T and published proof P, we could believe P was completely valid (i.e., each inference was valid) with probability 0.5 yet believe T was true with a probability of greater than 0.9.

In the previous sub-section, we argued that mathematicians may attempt to understand and increase their overall confidence in a proof by viewing the proof as a series of inferences and looking carefully at each inference within a proof. Figuratively speaking, we say they attempt to reach these goals by *zooming in* on the problematic parts of the proof. We conjecture that mathematicians also reach these goals by *zooming out* and looking at the high-level structure of the proof and thinking carefully about, not the individual inferences, but the *ideas* or *methods* in the proof.

Rav (1999) discusses how mathematicians do not focus on the logical details of the argument they are reading or even the theorem being proved. Rather, they look at the mathematical machinery being used to deduce new results from established ones. We believe that mathematicians read proofs to locate these methods, and that it is these methods that are useful to participants in their own professional work. Rav (1999) further argues that the reliability of proof does not stem from its *logical components*, but from its *methodological components* (p. 29). In addition to, or perhaps instead of, viewing proofs as a lengthy sequence of derivations, we hypothesize that mathematicians might encapsulate strings of derivations into a short collection of methods and determine whether these methods would allow one to deduce the claim that was proven. In other words, mathematicians viewed the proof as the application of a sequence of methods. Konoir's (1993) analysis of the structure of written proofs supports our claim. He found that proofs were often written with cues indicating to the reader how the proof should be partitioned and what methods were being applied in each partition. Similarly, in describing how he reads a proof in a field that he is familiar with, Thurston (1994) made the following remarks:

> I concentrate on the thoughts that are between the lines. I might look over several paragraphs or strings of equations and think to myself, 'oh yeah, they're putting in enough rigamorole to carry out such-and-such idea'. When the idea is clear, the formal set-up is usually unnecessary and redundant—I often feel that I could write it out myself more easily than figuring out what the authors actually wrote (p. 367).

M5 indicates that, due to time constraints, he sometimes behaves similarly. Several participants cited the benefits of breaking a proof into modules as a tool they used to understand proofs better. We believe that much of proof comprehension may consist of coordinating inspections of the logical details of a proof while zooming out to see these logical details in the context of the methods they are being used to support.

Our hypothesis that mathematicians validate proofs by examining whether the general methods in the proof would work rather than if each step in the proof was valid is also supported by Manin's (1977) and Hanna's (1991) claim that mathematicians evaluate a proof more by the plausibility of the argument in the proof than by its logical details and Thurston's (1994) claim that mathematicians can detect errors in a proof by thinking carefully about the ideas in the proof rather than validating its formal correctness. If this hypothesis is correct, it explains how theorems that appear in journals are usually true even if the proofs that accompany them are often flawed. High-level arguments could be correct despite logical flaws in the detail in the proof, and mathematicians are much better at evaluating the former than the latter; in fact, some may not spend much time on the latter task at all.

It is difficult to say how mathematicians "zoom out" from a proof by encapsulating particular strings of inferences of a proof into methods and then determine if those methods are valid. We suspect that this process gets at the heart of mathematical reasoning and is as cognitively complex as any task in mathematics. We imagine this process would have to involve structural-intuitive evidence, using one's intuition about the mathematics being discussed to see how the mathematical methods being used would work in one's mental models. Kreisel (1985) refers to the process of comparing a formal mathematical argument with one's mathematical background knowledge as "cross-checking" and Kreisel (1985), Thurston (1994), Otte (1994), and Rav (1999) argue that this process is essential in evaluating a proof.

A confirmatory quantitative study

As our model was generated by a relatively small number of mathematicians, we sought to obtain greater confidence in the generality of our results by adopting the methodology of Heinze (2010), who recommended complementing qualitative data and philosophical analyses with quantitative studies to build a more robust understanding of mathematical practice. Heinze constructed a survey to explore the different criteria that mathematicians employed to accept mathematical arguments. Our survey is similar, exploring why and how mathematicians read the published proofs of their colleagues.

Method

Following the methodology employed by Inglis and Mejia-Ramos (2009), we collected data through the internet in order to maximize our sample size. Recent studies have examined the validity of internet-based experiments by comparing a series of internet-based studies with their laboratory equivalents (e.g. Kranz & Dalal, 2000; Gosling et al., 2004). The notable degree of congruence between the two methodologies suggests that, by following simple guidelines, internet data has comparable validity to more traditional data. To ensure the validity of our data, we adopted the safeguards recommended by Reips (2000); in particular, we logged the IP address of each participant to screen for cases of multiple submissions by the same individual, we ensured the dropout rate was reasonable (in our case, under 25%), and ensured there were no statistical differences in the demographics or responses between those who completed our survey or left the study early. Given our adherence to Reips' (2000) guidelines, and the impracticality of obtaining large samples of research-active mathematicians in any other fashion, we believe our methods were justified.

We recruited mathematicians to participate in this study as follows. Twenty-four secretaries from top-ranked mathematics departments in the United States were contacted and asked to distribute an email to the mathematics faculty, post-doctoral researchers, and PhD students of that department. A total of 118 mathematicians agreed to participate. When participants clicked on the link to the survey website, they were taken to a webpage that described the purpose of the study and asked for demographic information, including their status (doctoral student, post-doc, or mathematics faculty). Of the 118 participants, 65 were doctoral students, 19 were post-docs, 33 were mathematics faculty, and one participant did not respond. Similar to Heinze (2010), post-hoc comparisons comparing the response patterns of mathematicians with different levels of experience revealed minimal differences (see Mejia-Ramos & Weber, submitted).

After completing the demographic information, participants were shown a screen saying they "will be asked about what you do when you are reading a proof that a colleague *published in a respected academic journal*" (the italics appeared in the text to participants). They were then asked to declare the extent to which they agreed (strongly disagreed, disagreed, neither agreed nor disagreed, agreed, or strongly agreed) with each one of 17 statements about why and how they read published proofs. The statements were based on hypotheses generated from our qualitative study and are presented in Table 1. Except for statements M2 and C1, all statements began with "When I read a proof in a respected journal." In general, the statements were of the form "it is not uncommon that [I engage in a hypothesized behavior]". We asked the question in this way as we did not want to see if participants *always* engaged in a behavior, but whether they engaged in the behavior *more than rarely*.

In the survey, we included three "foil" questions of behaviors that we did not think mathematicians would engage in (e.g., reading a proof to explore the writing styles of academics from different countries). These were included to verify that participants would not agree to saying it was not uncommon that they engaged in any plausible behavior.

Results

The results of the survey are summarized in Table 2. We included participants' mean response for each question, giving a +2 score if a participant strongly agreed with a statement, a +1 if the participant agreed, a 0 for a choice of "neither agreed nor disagreed", a -1 if the participant disagreed, and a -2 for strong disagreement with the given statement. As we were concerned about participants' responses to 17 items to assess statistical significance for an alpha-level of .05, we used a Bonferroni correction and set the alpha-level to .003. We first note the foils had their desired effect. For each foil, most participants did not agree with the foils and the majority of participants disagreed with them. This indicates most participants did not simply agree with any of these statements.

The hypotheses generated in Weber and Mejia-Ramos (2011) concerned participants reading the proofs of their colleagues. The data in Table 2 largely support each of these hypotheses.

Table 1. Questions used in the survey

Purposes:

P1: … it is not uncommon that I do not check the proof for correctness. Rather I read the proof to gain some other type of insight.

P2: … it is not uncommon that an important reason for my reading the proof is to gain some insights into how I can solve problems that I am working on.

Examples

E1: … it is not uncommon for me to see how the steps in the proof apply to a specific example. This increases my confidence that the proof is correct.

E2: … it is not uncommon for me to see how the steps in the proof apply to a specific example. This helps me understand the proof.

E3: … and I am not immediately sure that a statement in the proof is true, it is not uncommon for me to increase my confidence in the statement by checking it with one or more carefully chosen examples.

E4: … and I am not immediately sure that a statement in the proof is true, it is not uncommon for me to gain a sufficiently high level of confidence in the statement by checking it with one or more carefully chosen examples to assume the claim is correct and continue reading the proof.

Proof as application of methods:

M1: … it is not uncommon for me to try to understand the proof in terms of its main ideas and not only in terms of how each step is justified.

M2: When I read a lengthy proof in a respected journal, it is not uncommon that I skim the proof first to comprehend the main ideas of the proof, prior to reading the proof line-by-line.

M3: … it is not uncommon for me to judge the proof to be correct if I am sure the main idea or method is correct (as opposed to having to check every line of the proof).

M4: … it is not uncommon for me to try to understand the proof in terms of the overarching method the author used.

M5: … if I understand the main idea of the proof and think it is correct, it is not uncommon that I do not check that every line of the proof is correct, but trust that the logical details are correct.

Proof as cultural artifact:

C1: It is not uncommon for me to believe that a proof is correct because it is published in an academic journal.

C2: … it is not uncommon for the quality of the journal to increase my confidence that the proof is correct.

C3: … it is not uncommon for me to be very confident that the proof is correct because it was written by an authoritative source that I trust.

Foil:

F1: … it is not uncommon that I gain confidence that the proof is correct because the author cited my work.

F2: … it is not uncommon that an important reason that I read proofs is to explore the writing styles of academics from different countries.

F3: … it is not uncommon that I also read each of the references cited by the author.

Table 2. Participants' responses to the survey questions

Question	Mean	Agree	Neutral	Disagree
P1	0.82*	74%	14%	13%
P2	1.27*	90%	7%	3%
E1	0.97*	81%	8%	11%
E2	0.99*	82%	8%	9%
E3	0.86*	79%	13%	8%
E4	0.53*	56%	30%	14%
M1	1.31*	91%	3%	6%
M2	1.37*	92%	4%	3%
M3	0.78*	75%	15%	10%
M4	0.88*	77%	18%	5%
M5	0.88*	77%	14%	9%
C1	0.73*	72%	16%	12%
C2	0.63*	67%	16%	17%
C3	0.97*	83%	10%	7%
F1	-0.67*	6%	40%	54%
F2	-1.31*	5%	7%	88%
F3	-0.90*	11%	11%	78%

*- Indicates the mean was significantly different than zero with an alpha level of .003.

For all 14 items, the majority of participants agreed with our hypotheses about their mathematical practice, fewer than 20% disagreed, and their mean score was reliably greater than zero. For 12 of the 14 items, we had agreement levels of over 70%. In particular, these data support our hypotheses that mathematicians do not mainly read proofs in journals to check for correctness (P1), but often read proofs to gain insight into how to solve problems they were working on (P2). They use the cultural history of the proof, including who wrote the proof and the journal in which the proof appeared, to gain confidence that the proof was correct (C1, C2, C3). When analyzing the proof at a line-by-line level, mathematicians use examples, both to understand the proof (E2), to increase one's confidence that the proof is correct (E3), and sometimes as a primary means of verifying the validity of individual steps of the proof (E4). Mathematicians also try to understand the proof in terms of its overarching methods or big ideas

(M1, M4), using this to obtain a high degree of confidence in the validity of the proof (M3) and sometimes obviating the need to check each inference within the proof (M5).

Discussion

Summary of results.

In this paper, we presented a model for how mathematicians read the proofs of others, both to check for correctness and to gain insight. Mathematicians' confidence in a proof can be increased in three ways: (a) if the proof appears in a reputable source, some mathematicians will believe the proof is likely to be correct, (b) mathematicians will verify that each line in a proof validly follows from previous assertions, and (c) mathematicians may evaluate the overarching methods used in the proof and determine if they are appropriate, in addition to, or perhaps in lieu of, inspecting each line of the proof. Also, mathematicians attempt to understand proofs by studying not only (b) how each assertion in a proof follows from previous assertions, but also (c) what overarching methods in a proof could be useful to them to prove conjectures in their own work.

What is interesting is that each of these ways seems to rely, in part, on non-deductive evidence. For (a), trusting that a proof is reliable because it appears in a reputable source involves deferring to the authority of the editor of the journal and the reviewers who certified the proof as valid. While (b) can involve marshaling deductive evidence (i.e., the construction of sub-proofs), the data presented in Weber (2008), and supported by the data presented in this paper, show mathematicians sometimes rely on empirical evidence as well. We conjecture that (c) also involves the use of structural-intuitive evidence. Just as Boero (1999) observed that non-deductive argumentation plays a role in many of the stages of forming a conjecture and proof (although not in the actual written conjecture and proof), we argue that non-deductive argumentation plays an interesting role in the evaluation and comprehension of proof.

Caveats and limitations.

We supported our claims using a mixed methods approach, where our hypotheses were first generated and illustrated based on open-ended interviews with nine mathematicians about their professional practice reading proofs. We then assessed the viability of the hypotheses we generated with a survey distributed to 118 mathematicians. We note that there are still two challenges to the validity of our findings. The first is that the participants were self-selected. The participants for both the qualitative and quantitative study volunteered to participate; we think it is plausible that these mathematicians were more reflective about their practice than the typical mathematician and perhaps were more likely to engage in (or to admit to engaging in) the behaviors reported in this paper. Nonetheless, we do note for 12 of our 14 survey items, we had an agreement rate of over 70%, which does suggest the main effects we found from the survey are rather robust. A second limitation is that the data from both studies was self-report; we did not actually observe mathematicians engaging in these behaviors. It is possible that mathematicians' perceptions of their professional practice is not accurate (Inglis & Alcock, in press). This concern could be addressed with future task-based cognitive studies or sociological or ethnographic studies.

Implications for research and teaching.

Research in mathematics education on proof frequently consists of identifying common student behaviors with respect to conviction and proof that are at variance with mathematicians' treatment of conviction and proof. The model presented in this paper call into question some findings of this type. First, researchers commonly complain that believe a theorem is correct because it appears in a textbook (e.g., Harel & Sowder, 1998, p. 247). However, we wonder if this is really inconsistent with mathematical practice, as this paper presents strong evidence that many mathematicians are willing to accept that a theorem is correct because it appears in a journal. Second, researchers often lament that students will still check whether a theorem is true for examples, even after reading and accepting a general proof of the theorem, as Fischbein (1982) famously illustrated. Again, we argue this is actually consistent with mathematicians' behavior—as Mathematician 5 from the qualitative study noted, when he read a deductive proof, it was easy for him to overlook a mistake and believe something that as not true; checking the theorem or proof with specific examples was necessary to determine that everything made sense.

Perhaps what bothers mathematics educators with these behaviors is the beliefs we attribute to students who engage in them. Harel and Sowder (1998) contend students who rely on authority do so because they believe mathematics is a body of facts where the origin of those facts is inconsequential (p. 247) and Fischbein (1982) was concerned that students might not appreciate the generality of a deductive proof. Presumably mathematicians would not hold similar opinions. In this sense, the students and mathematicians *are* different. However, if it is these beliefs that mathematics educators find problematic, then this is what they should seek to change. Having students not rely on authority or examples to increase their confidence in ostensibly proven statements seems to us to be both unrealistic and undesirable.

More generally, we contend that many mathematics educators' treatment of different warrant-types is made on the basis of whether these types of evidence can provide absolute confidence that an assertion is correct or whether arguments based on these types of evidence would be sanctioned as proofs. Our perception is that, based on how some mathematics educators perceive mathematicians' practice, they view empirical and structural-intuitive evidence as useful for forming conjectures and suggesting methods of proof, but fundamentally insufficient in obtaining full conviction that a conjecture is correct. In contrast, a deductive argument can provide absolute certainty that a claim is correct and can therefore obviate the need for seeking other types of evidence. Authoritative evidence is dismissed as non-mathematical.

Table 3. Affordances and limitations of different warrant-types

Warrant Type	Level of Conviction	Potential for Insight	Socially acceptable for proof
Authoritative	Depends on the reputation of the source, possibly quite high	No insight	Not a proof
Naïve empiricism	Depends on whether empirical trends in the domain being studied are reliable, possibly quite high	No insight	Not a proof, but suggests conjectures
Generic proof	Depends on how transparent the chosen example is, possibly quite high	Illustrates why a claim is true	Sometimes accepted as an informal proof
Structural-intuitive	Depends on one's confidence in the consistency of his/her mental models and the formal theory, possibly quite high	Intuitive explanation for why a claim is true	Sometimes accepted as an informal proof (e.g., "proofs without words")
Deductive	Depends on the complexity of the proof and one's comfort with the proving methods being applied, possibly quite high but also not that high	Can highlight crucial properties or serve as basis for generic proof	Proof

We contend that such a perspective is too simplistic to account for the richness of mathematical practice. We give what we believe is a more comprehensive framework of the affordances of the different warrant-types in Table 3. We note several features of this table. First, we do not believe that any warrant-type provides *absolute* conviction; in fact, we think *complete* conviction in a mathematical assertion is rarely obtained by a single source of evidence in advanced mathematics. Second, we argue that proof, in particular, might not provide complete conviction because mathematicians may not be able to determine with certainty that the proof is correct, as some participants in our study suggested (See also Devlin, 2003, for contemporary examples of proofs of important conjectures that were accepted by the mathematical community before being shown to be irreparably flawed). Third, any warrant-type has the potential to provide high levels of conviction, where the level of conviction provided depends upon the context in which it is used. For instance, in Weber (2008), one mathematician indicated he would use naïve empiricism to verify a claim in a number theory proof as this type of reasoning tends to be reliable in modular arithmetic. However, he also added that he would not do so in topology, his area of research, since checking examples in a systematic order is more difficult. Fourth, we note that these types of evidence have a *social component*. Deductive evidence is preferred, at least in part, because arguments based upon it are socially sanctioned as mathematical proofs, not necessarily because this evidence provides higher levels of conviction than other arguments. In fact, some philosophers (e.g., Brown, 1997) argue that proofs based on pictures are as reliable, if not more reliable, than purely deductive arguments.

Most importantly, although all warrant-types are similar in that they can potentially provide either very high or modest levels of conviction, they differ with respect to the insight they can provide. Authoritative and naïve empiricism can, in some cases, convince mathematicians that a claim is true or an aspect of a proof is correct, but cannot explain why a claim is true or a proof is correct. This has an important implication for education. That students rely on authority as their source of conviction is not problematic because these students might come to believe things that are not true—indeed, given students' inability to validate proofs (Selden & Selden, 2003), they may be better off simply trusting the word of their teacher. The reliance on authority is harmful because it denies them the opportunity to gain insight from producing or studying a proof.

More generally, some view a goal of instruction as having students cease seeking conviction from authority or examples, but instead seek conviction by deductive reasoning (e.g., Harel & Sowder, 2007). We agree that students should become aware of the limitations of empirical, authoritative, and structural-intuitive warrants and they should appreciate the power and generality of deductive arguments. However, we disagree that students should *never* seek conviction through examples or authority or that they should no longer seek conviction in a theorem once a deductive proof has been produced. At least the data in this paper show that asking students to do this would be inconsistent with mathematicians' practice. Our view is that students should recognize the strengths and weaknesses of each warrant type, including the insight they can provide, the threats to their validity, and whether (and when) arguments based on this evidence would be socially sanctioned as convincing or proofs. As Harel (2001) showed, demonstrating how empirical arguments can be deceiving does little to change students' practice. Illustrating the insight provided by deductive arguments but not naïve empiricism does lead students to appreciate, and engage in the construction of, deductive proofs.

References

Alcock, L. & Weber, K. (2005). Proof validation in real analysis: Inferring and evaluating warrants. *Journal of Mathematical Behavior, 24*(2), 125-134.

Boero, P. (1999, July/August). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof.* Retrieved March 20, 2012, from http://www.lettredelapreuve.it/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html

Brown, J. (1997). Proofs and pictures. *British Journal for the Philosophy of Science*, 48, 161-180.

Davis, P. (1972). Fidelity in mathematical discourse: Is one and one really two? *American Mathematical Monthly, 79*, 252-263.

De Milo, R., Liptus, A., & Perlis, A. (1979). Social processes and proofs of theorems and programs. *Communications of the ACM, 22*, 271-280.

Devlin, K. (2003). The shame of it. From *Devlin's Angle*, an on-line publication of the MAA. Downloaded from: http://www.maa.org/devlin/devlin_archives.html. Last downloaded: March 20, 2012.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9-18.

Hanna, G. (1991). Mathematical proof. In D. Tall (ed.) *Advanced Mathematical Thinking.* Dordrecht: Kluwer.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM, 40*, 345-353.

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zazkis (Eds.) *Learning and teaching number theory*. New Jersey: Ablex Publishing Company.

Harel, G. & Sowder, L. (1998). Students proof schemes. *Research in Collegiate Mathematics Education, 3*, 234-282.

Harel, G. & Sowder, L. (2007). Towards a comprehensive perspective on proof. In F. Lester (Ed.), *Second handbook of research on mathematical teaching and learning*. Washington, DC: NCTM.

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

Heinze, A. (2010). Mathematicians' individual criteria for accepting theorems as proofs: An empirical approach. In G. Hanna, H.N. Jahnke, & H. Pulte (Eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives* (pp. 101-111). New York: Springer.

Inglis, M. & Mejia-Ramos, J.P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction, 27*, 25-50.

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modelling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66,* 3–21.

Jackson, A. (2006). Conjectures no more? Consensus forming on the proof of the Poincare and geometrization conjectures. *Notices of the American Mathematical Society, 53*, 897-901.

Konoir, J. (1993). Research into the construction of mathematical texts. *Educational Studies in Mathematics*, 24, 251-256.

Krantz, J. H., & Dalal, R. (2000). Validity of web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35–60). San Diego: Academic Press.

Kreisel, G. (1985). Mathematical logic: Tool and object lesson for science. *Synthese, 62*, 139-151.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Education Research Journal, 27*, 29-63.

Maher, C., Muter, E., & Kiczek, G. (2007). The development of proof making by students. In P. Boero (Ed.), *Theorems in Schools*. Rotterdam: Sense Publishers.

Manin, Y. (1977). *A course in Mathematical Logic*. New York: Springer Verlag.

Martin, W. G. & Harel, G. (1989). Proof frames of pre-service elementary teachers. *Journal for Research in Mathematics Education*, 20(1), 41-51.

Mejia-Ramos, J.P. & Weber, K. (submitted). How and why mathematicians read proofs: Further evidence.

RAND Mathematics Study Panel (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. Santa Monica, CA: RAND corporation.

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

Reips, U. D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 89–117). San Diego: Academic Press.

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.

Thurston, W.P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*, 161-177.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*, 431-459.

THE EVOLUTION OF CLASSROOM MATHEMATICAL PRACTICES IN A MATHEMATICS CONTENT COURSE FOR PROSPECTIVE ELEMENTARY TEACHERS

Ian Whitacre
San Diego State University

I report on the classroom mathematical practices that developed in a mathematics content course for prospective elementary teachers. The course focused on number and operations and was intended to promote number sense development. Instruction was guided by a local instruction theory for number sense development, which has been described previously. The present report focuses on the classroom mathematical practices that emerged and became established in the class during a recent teaching experiment. The actual learning route identified informs elaboration and refinement of the local instruction theory and sheds light on prospective teachers' number sense development.

Key words: Classroom mathematical practices, design research, local instruction theory, number sense, prospective teachers

I report on results of an analysis of collective activity (Rasmussen & Stephan, 2008) in a mathematics content course for prospective elementary teachers. In a previous study, Nickerson and I found that prospective elementary teachers involved in a classroom teaching experiment developed improved number sense, particularly in the form of flexible mental computation (Whitacre & Nickerson, 2006; Whitacre, 2007). Instruction in the previous teaching experiment was guided by a conjectured local instruction theory and informed the refinement and elaboration of that local instruction theory (Nickerson & Whitacre, 2010). The present study concerns a recent iteration of the classroom teaching experiment, in which the local instruction theory guided instructional planning. This report focuses on collective activity. I describe an actual learning route for prospective elementary teachers' number sense development in terms of classroom mathematical practices (Rasmussen & Stephan, 2008).

Background

I report new findings belonging to a program of research that addresses an important problem in undergraduate mathematics education: the mathematical preparation of prospective elementary teachers. Reform recommendations call for children to engage in meaningful mathematical activity, including making conjectures, developing their own solution strategies, making connections across mathematical topics, and participating in discussions of mathematical ideas (National Council of Teachers of Mathematics [NCTM], 1991; National Research Council [NRC], 2001). Reform visions of mathematics learning place added demands on teachers. Teachers need to be sensitive to their students' mathematical thinking and able to make sense of that thinking, and this in turn requires deep understanding of the mathematics itself (Jacobs, Lamb, & Philipp, 2010). However, prospective and practicing elementary teachers often know the procedures of elementary mathematics but do not understand the material conceptually (Ball, 1990; Ma, 1999; Thanheiser, 2010). *Even after* having taken their college mathematics courses, this population has been characterized as dependent upon the standard algorithms for elementary arithmetic (Ball, 1990; Newton, 2008; Yang, 2007). Because we depend upon elementary teachers to educate children, their mathematical preparation is an important concern.

Number Sense & Mental Computation

I find it useful to conceptualize the problem of prospective elementary teachers' insufficient mathematical preparation in terms of number sense. Reys and Yang (1998) describe *number sense* as follows:

> Number sense refers to a person's general understanding of number and operations. It also includes the ability and inclination to use this understanding in flexible ways to make mathematical judgments and to develop useful strategies for handling numbers and operations… (pp. 225–226)

Number sense is recognized as an important goal of mathematics instruction (NCTM, 2000; NRC, 2001). However, children both within the United States and internationally tend to learn mathematics in a way that emphasizes the rote application of standard algorithms and does not support their development of number sense (Reys, Reys, McIntosh, Emanuelsson, Johansson, & Yang, 1999). In order to ameliorate this situation, mathematics educators have a responsibility to positively influence the number sense of prospective elementary teachers.

Number sense is not a mathematical topic per se. It is a characterization of how individuals reason mathematically. However, certain mathematical tasks lend themselves to the exercise of number sense moreso than others. In order for flexibility to be a possibility, the nature of the task has to allow for various ways of reasoning. In the mathematics education literature, three topics that have been associated with number sense are mental computation (Heirdsfield & Cooper, 2004; Markovits & Sowder, 1994), computational estimation (Sowder, 1992), and reasoning about fraction magnitude (Whitacre & Nickerson, 2011; Yang, Reys, & Reys, 2009). Number sense is related to mental computation in terms of the flexibility with which a person approaches mental computation, as well as the tendency to use nonstandard strategies (Heirdsfield & Cooper, 2004; Markovits & Sowder, 1994; Whitacre, 2007).

Heirdsfield and Cooper (2004) describe the processes of inflexible and flexible mental calculators.[1] For inflexible mental calculators, operations map to particular algorithms. Their way of performing an operation mentally is simply to use the mental analogue of the standard algorithm for that operation. Flexible mental calculators, by contrast, make a choice of strategy that is sensitive to the numbers involved in the computation and informed by knowledge of numeration and number facts, understand of the effect of operations on numbers, knowledge of strategies, and beliefs about strategies. Thus, when choices are made, the person's number sense is exercised. Engagement in mental computation can lead to improved number sense (Sowder, 1992; Whitacre & Nickerson, 2006). Furthermore, the choices that a person makes in mental computation provide a window into her number sense (Heirdsfield & Cooper, 2004; Markovits & Sowder, 1994).

Markovits and Sowder (1994) relate whole-number mental computation strategies to number sense according to the extent to which strategies depart from the standard written algorithms. Mental computation strategies are grouped into the categories of Standard, Transition, Nonstandard with no reformulation, and Nonstandard with reformulation. The less similar to the mental analogue of the standard algorithm (MASA) a strategy is, the more indicative it is of good number sense. This connection hinges on the idea that individuals who use a variety of nonstandard strategies make a choice of strategy depending on the particular numbers involved

[1] Heirdsfield and Cooper studied the processes of accurate flexible and accurate inflexible mental calculators. The focus here is the distinction between flexibility and inflexibility.

in a computation and that when a person chooses to use a nonstandard strategy, this tends to be one that makes sense to the person using it.

The Standard-to-Nonstandard framework influences the conceptualization of number sense development reflected in this study. At the same time, in the course of this research, I have found the need to modify the definitions of the categories. For example, Markovits and Sowder defined *Nonstandard with no reformulation* specifically in terms of the student using a “left-to-right process.” However, there are strategies that prospective elementary teachers use, namely Aggregation strategies, that are nonstandard and do not involve reformulation but are not accurately characterized as left-to-right processes. I view both right-to-left and left-to-right processes as Transition strategies because both involve separating numbers into tens and ones and computing place-value-wise, rather than working with the numbers as whole amounts. Furthermore, students who use one of these often use the other. They may choose between a right-to-left or left-to-right process depending on whether or not regrouping is necessary. Thus, for prospective elementary teachers, these strategies seem to be closely related. The revised definitions of the categories are in keeping with the spirit of Markovits and Sowder’s framework. Furthermore, ordering of specific strategies from Standard to Nonstandard is unaffected by the revisions to the definitions.

Since this paper focuses on addition and subtraction activity, I describe the categories specifically in relation to reasoning about addition and subtraction:

Standard: Using the mental analogue of the standard addition or subtraction algorithm

Transition: Using a right-to-left or left-to-right process

Nonstandard with no reformulation: Beginning with one of the given numbers and increasing or decreasing according to the other

Nonstandard with reformulation: Rounding one or both numbers, computing, and then compensating if necessary

Table 1 presents specific mental addition and subtraction strategies belonging to each category.

Table 1. Prospective Elementary Teachers’ Standard-to-Nonstandard Mental Addition and Subtraction Strategies

	Standard	Transition	Nonstandard with no reform.	Nonstadard with reformulation
Addition	MASA	Right to Left Left to Right	Aggregation	Giving[2] Single Compensation Double Compensation
Subtraction	MASA	Right to Left Left to Right	Aggregation	Minuend Compensation Subtrahend Compensation Shifting the Difference

Theoretical Perspective

The emergent perspective (Cobb & Yackel, 1996) informs my thiking about the phenomenon of prospective elementary teachers’ number sense development. This perspective represents a

[2] Heirdsfield and Cooper (2004) call this Levelling. I use the term Giving because prospective elementary teachers often talk about this strategy in terms of “giving” part of one number to the other number.

coordination of sociocultural and constructivist approaches. In particular, I view social norms, sociomathematical norms, and classroom mathematical practices as reflexively related to their individual correlates. (See Fig. 1.) From this perspective, classroom mathematics learning occurs in the doing of activities within a culture. The nature of those activities and the more general classroom culture profoundly shape what is learned. At the same time, the classroom culture is created in interactions between individuals participating in classroom activities. In fact, without them, no classroom culture exists. Classroom members' ways of participating in activities are influenced by their beliefs, values, and conceptions. Our previous empirical reports have focused on change in individuals' mathematical conceptions and activity (Whitacre & Nickerson, 2006; Whitacre, 2007). This report focuses on classroom mathematical practices.

Social Perspective	**Psychological Perspective**
Classroom social norms	Beliefs about own role, others' roles, and the general nature of mathematical activity in school
Sociomathematical norms	Mathematical beliefs and values
Classroom mathematical practices	Mathematical conceptions and activity

Fig. 1. Interpretive framework (Cobb & Yackel, 1996)

Local Instruction Theory

The study reported here is part of an ongoing design research effort (Cobb & Bowers, 1999), which focuses on prospective elementary teachers' number sense development. This research program takes the form of both classroom teaching experiments and theory building, and these are reflexively related. Nickerson and I have developed a local instruction theory for number sense development, which continues to evolve as our research progresses. A *local instruction theory* (LIT) refers to "the description of, and rationale for, the envisioned learning route as it relates to a set of instructional activities for a specific topic" (Gravemeijer, 2004, p. 107). In a recent publication, we have described in some detail our LIT for number sense development (Nickerson & Whitacre, 2010). In that report, we separate the LIT into three specific goals and present accompanying envisioned learning routes. That presentation is useful for communicating clearly and independently about three distinct aspects of the activity—strategies, discourse, and models. In this report, I describe the LIT briefly and in unified terms.

Essentially, the goal is for prospective elementary teachers to move from dependence on standard algorithms to reasoning flexibly about numbers and operations. I conceptualize this process in terms of movement along the spectrum from Standard to Nonstandard. This is not to say that students necessarily leave behind Standard or Transition strategies as they make sense of and use nonstandard strategies. Rather, development looks like students broadening their repertoires of strategies and thus becoming more able to choose a suitable strategy for a given computation.

Previous Research

This design research program has proceeded through two cyles of instructional design, classroom teaching experiments, data analysis, and theory building. In the first classroom teaching experiment, instruction was guided by a conjectured local instruction theory. At that time, it was a question whether incorporating mental computation as an authentic activity (Brown, Collins, & Duguid, 1989) in the content course would be a viable approach to facilitating students' number sense development. I analyzed change in students' number sense using pre/post interviews with 13 participants. Given a range of mental addition, subtraction, and multiplication problems in story contexts, the participants became more flexible and shifted from favoring the mental analogues of the standard algorithms to favoring nonstandard strategies (Whitacre, 2007). These results, together with our impressions from classroom activity, were encouraging. They also informed revisions and elaboration to the local instruction theory. In a recent classroom teaching experiment, the revised LIT guided instruction. Seven of the students were interviewed pre and post, and I found similar improvements in their number sense. They too became more flexible in mental computation and came to favor nonstandard strategies (Whitacre, 2012).

In this recent classroom teaching experiment, I asked new research questions.

Methods

The research question that is the focus of this report is the following: In a mathematics content course for prospective elementary teachers, which is guided by a local instruction theory for the development of number sense,

What classroom mathematical practices emerge and become established?

I view the progression through classroom mathematical practices as an actual learning route toward number sense development. In the spirit of design research, this actual learning route facilitates an improved understanding of the phenomenon and elaboration and refinement to the local instruction theory.

Data collection took place during fall of 2010 in a mathematics content course taught at a large, urban university in the southwestern United States. There were 39 students enrolled in the course, and 38 of the students were female. The majority of the students were freshmen Liberal Studies majors. The instructor of the course was Dr. Susan Nickerson. She is a mathematics educator and an experienced teacher of mathematics courses for prospective teachers. The data corpus for the study reported here consisted of videotapes of 7 days of class. These were Days #3, 6, 7, 8, 9, 11, and 12 of the semester-long course. The class met twice weekly for 75 minutes per meeting. Three cameras were used to record classroom activity throughout the semester.

The methodology of Rasmussen and Stephan (2008) was used to analyze collective activity. This methodology involves coding arguments using Toulmin's (1969) model, which describes the anatomy of an argument in terms of claim, data, warrant, and backing. The claim is the assertion being made. The data is evidence offered in support of the claim. The warrant explains how the data supports the claim. Backing serves to justify the validity of the warrant.

The methodology is a three-phase process: (1) Whole-class discussions are transcribed. Researchers watch video of each discussion and identify any claims that are made. For each claim, an argumentation scheme is constructed, which explicitly identifies each of the components of the argument[3]. This analysis yields a chronological argumentation log. (2)

---

[3] Specifically, these are mathematical arguments. Non-mathematical arguments are not included in the analysis. However, what counts as mathematical depends on the course topic and on the researchers' interests. For example, in the context of the mathematics content course, claims concerning how to name mental computation strategies are considered mathematical. Also note claims that are both dataless and warrantless are ignored.

Reseachers look across the argumentation log to identify ideas that *functioned as if shared* in whole-class discussion. Criteria for ideas functioning as if shared are (i) warrants or backings dropping off, (ii) an element of an argument shifting roles (e.g., from claim to warrant), and (iii) repeated use of data or warrants in support of different claims (Cole, Becker, Towns, Sweeney, Wawro, & Ramussen, 2011). (3) The as-if shared ideas are then organized according to related mathematical activities to describe classroom mathematical practices.

Rasmussen and Stephan (2008) define a *classroom mathematical practice* (CMP) as a "collection of as-if shared ideas that are integral to the development of a more general mathematical activity" (p. 201). This definition differs from that of Cobb and Yackel (1996) in that a CMP is defined in terms of a set of mathematical ideas, rather than a single idea. This distinction has more theoretical significance than might be apparent at first glance. For Stephan, Cobb and Gravemeijer (2003), who did earlier work on classroom mathematical practices, the construct corresponded to a single taken-as-shared idea. Essentially, a sequence of taken-as-shared ideas was documented. By contrast, when a CMP is defined in terms of a *set* of as-if shared ideas, it is up to the researcher to organize those as-if shared ideas. The way in which as-if shared ideas are organized depends on the researcher's focus and sensitivity to phenomena of interest. In the case of the analysis reported here, a focus on number sense development informed the grain size and categorization criteria by which CMPs were identified.

This report focuses on two related strands of activity. Whole-class discussions concerning ideas of whole-number composition and place value occurred on Days 6, 7, and 8. A total of 74 arguments were made during this strand of activity. Whole-class discussions directly related to addition and subtraction occurred on Days 3, 8, 9, 11, and 12. A total of 48 arguments belonged to this strand. Occurrences of ideas belonging to either strand were considered in order to apply the criteria for ideas functioning as if shared.

The methodology of Rasmussen and Stephan (2008) was developed in the context of documenting collective activity in inquiry-oriented differential equations (Stephan & Rasmussen, 2002). It has also been used to analyze activity in inquiry-oriented linear algebra (Wawro, 2011) and chemistry (Cole et al., 2011). As such, it was unclear at the outset whether this approach would be appropriate to the analysis of collective activity in an elementary mathematics content course. I will revisit this point in the Discussion section.

Results

This section presents the CMPs belonging to two related strands of activity. They are presented in a unified manner. I present these chronologically (according the point in time at which a CMP became established) and attempt thereby to tell the story of the actual learning route that was traversed by the class. The succession of CMPs around number composition, place value, addition, and subtraction was as follows:

CMP1. Relying on standard algorithms and notation
CMP2. Making sense of additive composition and place value
CMP3. Making sense of Standard and Transition strategies
CMP4. Reasoning flexibly about addition
CMP5. Reasoning flexibly about subtraction

Below, I briefly describe the collective activity that characterized each CMP.

CMP1. Relying on standard algorithms and notation

In discussions of mental calculative work early in the course, the class behaved as if the authority of the standard algorithms was assumed. Mental computations using the mental analogues of the standard algorithms went unquestioned, whereas nonstandard strategies required mathematical justification. Early written records of mental computations were numeric-algorithmic in nature, with digits arranged in rows and columns, even when nonstandard strategies were used. Gesturing associated with the articulation of these strategies involved finger tracing up and down columns, essentially reenacting written work.

CMP2. Making sense of additive composition and place value

Students participated in a variety of activities involving counting, grouping, and regrouping. These activities concerned story problems involving quantities (apples), physical grouping of multilink cubes, and operating with numerals. The class transitioned from solving problems by means of drawings or through manipulation of physical objects to using place-value numeration systems to record numbers as numerals in various bases. Numerals were intrepreted in terms of groups according to the base, especially base eight, three, or ten. Many of the as-if shared ideas were specific to the context and/or base. These included how apples were packaged on Andrew's Apple Farm, including the specific numbers of apples that filled a basket, bushel, or truck; Andrew's method of bookkeeping, which involved an informal version of base-eight notation; grouping multilink cubes by threes, nines, and twenty-sevens; recording numerals in base three and interpreting digits as groups of a certain size; and using and interpreting place-value notation in base ten and other bases to represent the same numbers of items by converting between bases.

CMP3. Making sense of Standard and Transition strategies

Building on CMP2, in CMP3, students combined place-value ideas with the operations of addition and subtraction. Addition came to involve an aggregating and regrouping process, grounded in counting in the given base. The addends and sum were recorded as numerals in the given base. Regrouping moves were notated by writing a *1* above the digit in the next place to the left. Likewise with subtraction, "borrowing" took on the meaning of unpacking a group of size b. The minuend, subtrahend, and difference were recorded as numerals in the given base. Regrouping moves were notated in one of two ways: (1) by writing *1* to the left of the digit of the minuend in the place that received the extra items, or (2) by writing *10* above that digit. The as-if shared ideas that characterized CMP3 were separating numbers into tens and ones, regrouping from right to left, regrouping in order to subtract, reasoning about subtraction as a take-away process, and using and interpreting place-value notation in bases three and ten.

CMP4. Reasoning Flexibly about Addition

In CMP4, nonstandard addition strategies were used, discussed, justified, named, and compared. Students' justifications were grounded as-if shared ideas related to addition, namely, reasoning about addition as a cumulative process of increase, reasoning in terms of noncanonical number composition and decomposing numbers as convenient, and using rounding, reasoning about the effects of rounding, and compensating if necessary. The strategy that I call Giving became established for the class, and students named it "Borrow to Build."

CMP5. Reasoning Flexibly about Subtraction

In CMP5, students used nonstandard subtraction strategies to compute differences. Reasoning about subtraction was often closely related to the empty number line. The as-if shared used in students' arguments were reasoning about differences in terms of the distance between number-locations, reasoning about movement along the number line, reasoning about subtraction as a cumulative process of decrease, and reasoning about subtraction as a take-away process.

The validity of "Shifting the Difference" was established on the basis of maintaining the distance between these number-locations. Below, I present one of the arguments made by a student concerning "Shifting the Difference."

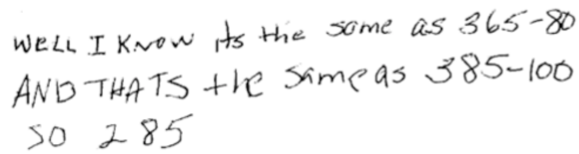


Fig. 1. Child's strategy for 364 – 79 (Sowder, Sowder, & Nickerson, 2010, p. 55)

On Day 12, students were asked to interpret examples of children's reasoning about the computation 364 – 79. One child's reasoning was represented by the written work in Fig. 1. Three students—Trina, Valerie, and Amelia[4]—made arguments for the validity of this child's strategy. Valerie's argument involved reasoning about difference as distance between:

Valerie: Okay, so we thought about it in terms of, when you're subtracting, you're trying to find the distance between two numbers. So, we thought of it kind of in terms of a number line… So, you started off with 79 and 364. So, 364 moved up one to 365 and also, likewise the 79 moved up to 80. So, the distance didn't change between the numbers. So, originally it was right here, and they both moved up one on a number line. So, the distance between them is the same. So, similarly when you have 385 and 100, you just added 21. So, if you took the numbers from their original position and moved them each up 21 spaces, the shift would be the same and the distance between both numbers is the same.

[Valerie uses her hands as number-locations. She moves both hands to her right as she talks about the numbers "moving up."]

Fig. Z. Valerie's gesturing during her argument

In Valerie's argument, reasoning about the difference as a distance between number-locations served as backing for the warrant that adding the same amount to both the minuend and subtrahend maintained the difference. In previous arguments, the idea of the difference as distance between had been used as data. The shift in its argumentative role coincided with and afforded the justification of a new idea: Shifting the Difference. Shifting the Difference, in turn, was used to justify the equal-additions algorithm. Empty-number-line inscriptions were integral to this classroom math practice, as was gesturing that illustrated distances spanned and shifted.

During the emergence and establishment of CMPs 4 and 5, a variety of nonstandard mental computation strategies became normative ways of computing sums and differences. Once students had made sense of the standard algorithms, they moved beyond them, using increasingly sophisticated addition and subtraction strategies. The ways of reasoning that characterized CMPs

---

4 Student names are pseudonyms.

4 and 5 are unusual for prospective elementary teachers, who typically rely heavily on the standard algorithms (Ball, 1990; Yang, 2007). These strategies are regarded as indicative of good number sense (Markovits & Sowder, 1994; Yang, Reys, & Reys, 2009).

Discussion

It is worth reminding the reader that the progression through classroom mathematical practices is based on arguments articulated in whole-class discussion and, as such, represents an *actual learning route* (analogous to an actual learning trajectory). Figure 2 relates the CMPs to the Standard-to-Nonstandard spectrum and, thereby, to the envisioned learning route. We find that the actual learning route followed the envisioned learning route, meaning that the chronological order of establishment of CMPs corresponded to the envisioned progression from Standard to Nonstandard strategies.

CMP1	CMP2 & CMP3	CMP4 & CMP5	
Standard	Transition	Nonstandard w/o Reformulation	Nonstandard with Reformulation

Fig. 2. Correspondence between CMPs and Standard-to-Nonstandard spectrum

Below, we discuss what we learn about number sense development from the results of this analysis. We also discuss the nature of classroom mathematical practices and address both practical and methodological questions related to documenting collective activity.

Revisiting a Local Instruction Theory for Number Sense Development

In the paper in which we described our LIT for number sense development (Nickerson & Whitacre, 2010), we focused on students' activities of naming strategies and using tools like the empty number line. The previous emphasis reflected a focus on the instructional design aspects of our work. In service of Goal 1, we sought to move prospective elementary teachers from dependence on the standard algorithms to capitalizing on opportunities to use number-sensible strategies. We knew from analyses of interview data that individual students had changed from dependence on the standard algorithms to reasoning flexibly about operations. However, we had not systematically analyzed collective activity to identify normative ways of reasoning.

The results presented here reflect a focus on the instructional sequence *as enacted* and highlight the development of students' ways of reasoning. The progression of classroom math practices that our analysis revealed complements the previous description of the LIT. It enables us to relate the envisioned learning route described in the LIT to an actual learning route, with students' activities providing the bridge between these.

Several aspects of the succession of CMPs illuminate our understanding of the phenomenon of prospective elementary teachers' number sense development. CMP1 corresponds to the Standard category. The class reasoned about the operations in a way that relied on, and assumed the authority of, the standard algorithms. In earlier presentations of the local instruction theory, we have described our students as beginning the course dependent on the standard algorithms (Nickerson & Whitacre, 2008; Nickerson & Whitacre, 2010). This expectation was based on a review of the literature concerning prospective elementary teachers' mathematical thinking, as well as our experience teaching the course. We conceptualized this expectation in terms of the reasoning of individual students, not of the collective. Indeed, we know from interviews with the participants in our teaching experiments that they tend to be dependent on the standard algorithms at the outset of the course. However, it does not follow from individual students' dependence on the standard algorithms that relying on standard algorithms would be a normative way of reasoning for the class. The analysis of collective acitivity revealed that initially standard algorithms functioned as authoritative. In a vacuum, this might appear to be an undesirable situation. However, the authority of the standard algorithms was leveraged productively, particularly through the instructor's skillful discursive moves (Rasmussen, Kwon, & Marrongelle, 2008) to motivate the need to justify nonstandard strategies.

Another unexpected result of the analysis concerned the idea of subtraction as a take-away process. Prospective elementary teachers come into the course reasoning about subtracting in terms of taking away; this is not a new idea to them. Furthermore, there are limitations to the take-away meaning for subtraction, and the course curriculum includes tasks to engage students in thinking about various meanings for subtraction in relation to story problems (Nickerson & Whitacre, 2007; Sowder, Sowder, & Nickerson, 2010). Thus, in a vacuum, it would not seem particularly desirable for reasoning about subtraction as a take-away process to become normative. However, this idea was used productively in the progression toward CMP5—Reasoning Flexibly about Subtraction. In particular, it was used to justify aggregation strategies for subtraction. As with assuming the authority of the standard algorithms, an idea that students brought with them to the course came to function as-if shared and served to advance the mathematical activity.

A third illuminating result of the analysis concerns the sequential relationship between CMP3 and CMPs 4 and 5. The class moved from making sense of the standard algorithms to reasoning flexibly about addition and subtraction. It might be expected that students' reasoning about number composition and place value would lead to their ability to make sense of nonstandard strategies. However, the ideas that came into play in justifying nonstandard addition and subtraction strategies actually had little to do with canonical number composition (in terms of ones, tens, and hundreds). Note the lack of overlap between the as-if shared ideas related to place value and number composition in CMPs 2 and 3 and the as-if shared ideas belonging to CMPs 4 and 5. This progression seems perplexing: CMPs 4 and 5 succeeded CMP3 chronologically but with little overlap structurally.

If those ideas did not clearly function to help the class make sense of nonstandard strategies, then what purpose did they serve? We believe it is not by coincidence that the ways of reasoning belonging to CMPs 4 and 5 became normative shortly after CMP3 was established. We conjecture that students' initial dependence on the standard algorithms explains this result. Thinking in terms of an individual prospective elementary teacher, she enters the course dependent on the standard algorithms. She knows that these algorithms are endorsed by authorities (teachers and textbooks); they are valid ways of computing. Yet, she does not understand why they work. She is not dependent on these algorithms despite her lack of understanding. She is dependent on these algorithms *because* she does not understand them.

The normative ways of reasoning in CMPs 2 and 3 relate directly to students' abilities to explain and justify the standard addition and subtraction algorithms. That is, these ideas represent ways of making sense of those algorithms. Once these algorithms make sense in terms of canonical number composition and place-value notation, then the fact that they are endorsed by authorities becomes less relevant. Students no longer need to depend on endorsed algorithms because their ability make sense of those algorithms also affords making sense of Transition strategies. Essentially, students move from viewing the algorithms in terms of digits that live in columns to viewing them in terms of numbers of ones, tens, and hundreds (Thanheiser, 2010). Then operating from left-to-right, rather than right-to-left becomes just as sensible. Once students are using more than one strategy for a given operation, the possibility of using a wider variety of strategies is not far behind. An essential criterion shift has occurred from justification based on endorsement or convention to justification based on mathematical validity. This is precisely the distinction between CMP1 and CMP3. The class moved from assuming the authority of the standard algorithms and questioning nonstandard strategies to reasoning on a valid mathematical foundation about ways of performing addition and subtraction.

Reflections on Documenting Collective Activity

The methodology employed in this study was developed around analyses of inquiry-oriented differential equations classes (Stephan & Rasmussen, 2002; Rasmussen & Stephan 2008). It has also been used to document collective activity in inquiry-oriented linear algebra (Wawro, 2011) and in physical chemistry (Cole et al., 2011). It was not clear whether this methodology would be appropriate to the study of an elementary mathematics content course. We found that it was a viable methodology to use with our data set. Furthermore, the results of the analysis illuminated our understanding of the phenomenon of interest.

The methodology served our purposes because the course that we studied was characterized by students' engagement in mathematical argumentation. I agree with Philipp (2008) that, "Elementary mathematics is not elementary" (p. 19). In particular, it is possible for students to engage in quite sophisticated mathematical activity while dealing only with whole numbers and basic arithmetic operations. As an illustration of this point, Wawro (2012) reported that she needed to use expanded argumentation schemes to document collective activity in inquiry-oriented linear algebra. She found that 22 of the 118 arguments in her data set had expanded structures. Wawro attributes the need for these expanded schemes to the complexity of students' arguments in a class in which they are transition to formal proof. In our analysis, we found that 20 of 122 arguments had expanded structures. Thus, even though the mathematics was elementary, the level of students' engagement with that mathematics was rather advanced.

Stephan and Rasmussen (2002) found that CMPs could be non-sequential in both time and structure. That is, they could co-exist temporally, and they could consist of overlapping as-if shared ideas. In the analysis reported here, we found a set of CMPs that were not strictly sequential in time or structure. CMPs 2 and 3 overlapped in time, as did CMPs 4 and 5. We note that, in part, this is a question of grain size and focus. It is up to the researchers' discretion to identify classroom mathemtical practices by categorizing as-if shared ideas as being related to some more general mathematical activity. For example, I could have taken a coarse grain size and grouped CMPs 2 and 3 and CMPs 4 and 5 together. I could also slice the activity differently. I distinguished CMPs 4 and 5 on the basis of the operation about which students were reasoning—addition versus subtraction. We could otherwise have sliced these on the basis of categories of strategies. This sort of decision influences whether the CMPs identified overlap in structure. My ultimate decision to distinguish operation was based on the fact that very different sets of ideas arose in students' arguments concerning addition versus subtraction.

Significance

The motivation for our research program stems from the troubling reality that prospective elementary teachers in the United States and elsewhere tend to be poorly prepared to teach mathematics effectively (Ball, 1990; Ma, 1999; Newton, 2008; Tsao, 2005; Yang, Reys, & Reys, 2009). In mathematics content courses like the one that we studied, mathematics educators have the opportunity to facilitate prospective teachers' number sense development and thus help them become better prepared to foster children's learning of mathematics. For this reason, analyses that illuminate processes by which prospective teachers develop improved number sense are valuable to the field. This research is also of methodological interest to the undergraduate mathematics education community. Researchers in the community are interested in instructional design theory and in analysis of collective activity in undergraduate mathematics classrooms.

References

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal, 90,* 449-466.

Brown, J. S., Collins, A., Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18,* 32-42.

Cobb, P., & Bowers, J. (1999). Cognitive and situated learning perspectives. *Educational Researcher, 28* (2), 4-15.

Cole, R., Becker, N., Towns, M., Sweeney, G., Wawro, M., Rasmussen, C. (2011). Adapting a methodology from mathematics education research to chemistry education research: Documenting collective activity. *International Journal of Science and Mathematics Education.* Retrieved on September 22, 2011, from: http://www.springerlink.com/content/m37642217677168m/

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

Gravemeijer, K. (2004). Local instruction theories as a means of support for teachers in reform mathematics education. *Mathematical Thinking and Learning, 6*, 105–128.

Greeno, J. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education, 22*, 170–218.

Heirdsfield, A. M., & Cooper T. J. (2004). Factors affecting the process of proficient mental addition and subtraction: case studies of flexible and inflexible computers. *Journal of Mathematical Behavior, 23,* 443–463.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States.* New Jersey: Erlbaum.

Markovits, Z. & Sowder, J. (1994). Developing number sense: An intervention study in grade 7. *Journal for Research in Mathematics Education, 25*, 4–29.

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

National Research Council (2001). *Adding it up: Helping children learn mathematics.* In J. Kilpatrick, J. Swafford, & B. Findel (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Newton, K. J. (2008). An extensive analysis of preservice elementary teachers' knowledge of fractions. *American Educational Research Journal, 45,* 1080-1110.

Nickerson, S. D., & Whitacre, I. (2010). A local instruction theory for the development of number sense. *Mathematical Thinking and Learning, 12*, 227-252.

Rasmussen, C., Kwon, O. N., & Marrongelle, K. (2008). A framework for interpreting inquiry-oriented teaching. *Proceedings of the Eleventh Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education.*

Rasmussen, C., Kwon, O.N., and Marrongelle, K. (2009). A Framework for Interpreting Inquiry-Oriented Teaching: Opportunities for Student and Teacher Learning. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Rasmussen, C., & Stephan, M. (2008). A methodology for documenting collective activity. In A. E. Kelly & R. Lesh (Eds.), *Design research in education* (pp. 195-215). Mawah, NJ: Erlbaum.
Sowder, J. (1992). Estimation and number sense. In D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 371–389). New York: Macmillan.
Tsao, Y.-L. (2005). The number sense of preservice elementary school teachers. *College Student Journal, 39,* 647-679.
Wertsch, J. (1991) Voices of the mind: A sociocultural approach to mediated action. Harvard University Press.
Whitacre, I., & Nickerson, S. D. (2006). Pedagogy that makes (number) sense: a classroom teaching experiment around mental math. In S. Alatorre, J. L. Cortina, M. Sáiz, & A. Méndez, (Eds.), *Proceedings of the twenty-eighth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics* Education, Vol. 2 (pp. 736-743). Mérida, México: Universidad Pedagógica Nacional.
Whitacre, I. (2007). Preservice teachers' number sensible mental computation strategies. *Proceedings of the Tenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education.* San Diego, CA. http://sigmaa.maa.org/rume/crume2007/papers/whitacre.pdf
Yang, D. C. (2007). Investigating the strategies used by preservice teachers in Taiwan when responding to number sense questions. *School Science and Mathematics, 107*, 293–301.
Yang, D. C., Reys, R. E., & Reys, B. J. (2009). Number sense strategies used by pre-service teachers in Taiwan. *International Journal of Science and Mathematics Education, 7*, 383–403.

STUDENT CONCEPT IMAGES OF FUNCTION AND LINEAR TRANSFORMATION

Michelle Zandieh
Arizona State University

Jessica Ellis
San Diego State University

Chris Rasmussen
San Diego State University

As part of a larger study of student understanding of concepts in linear algebra, we interviewed 10 university linear algebra students as to their conceptions of functions from high school algebra and linear transformation from their study of linear algebra. Analysis of these results led to a classification of student responses into properties, computations and a series of five interrelated clusters of metaphorical expressions. We see this classification as providing richness and nuance to existing literature on students' conceptions of function. In addition, we are finding these categories helpful in describing the compatibilities and distinctions in student understanding of function and linear transformation.

Keywords: Concept image, function, linear algebra, linear transformation, metaphor

The research reported in this paper began as part of a larger study into the teaching and learning of linear algebra. As we examined student understanding of linear transformations we wondered how student understanding of functions from their study of precalculus and calculus might influence their understanding of linear transformations and vice versa. In order to explore this issue, we found that we needed ways to describe student understanding of functions and linear transformations that might go beyond traditional characterizations of functions from the research literature. This proposal elaborates our new characterization and provides an example of how this characterization can be used to compare student understanding of function and linear transformation.

Literature and theoretical background

The nature of students' conceptions of function has a long history in the mathematics education research literature. This work includes Monk's (1992) pointwise versus across-time distinction, the APOS (action, process, object, scheme) view of function (e.g., Breidenbach, Dubinsky, Hawkes, & Nichols,1992; Dubinsky & McDonald, 2001), and Sfard's (1991, 1992) structural and operational conceptions of function. A comparison of these views may be found within Zandieh (2000). More recent work has focused on descriptions of function as covariational reasoning (e.g., Thompson, 1995; Carlson, Jacobs, Coe, Larsen & Hsu, 2002). A recent summary with a focus towards covariational reasoning is found in Oehrtman, Carlson, and Thompson (2008).

The work in linear algebra has tended to focus more on student difficulties (e.g., Carlson, 1993; Dorier, Robert, Robinet & Rogalski, 2000; Harel, 1989; Hillel, 2000; Sierpinska, 2000). There have been a few studies on student understanding of linear transformation (Dreyfus, Hillel, & Sierpinska, 1998; Portnoy, Grundmeier, & Graham, 2006). However, we could not find studies that relate student understanding of function and linear transformation.

In addition to work specifically on student conceptions of functions or linear transformation, we were interested in research that explores how one may characterize the conceptions that a student has for a particular mathematical construct. The term *concept image* has been used to refer to the "set of all mental pictures associated in the students' mind with the concept name, together with all the properties characterizing them" (Vinner & Dreyfus, 1989, p. 356). Tall and

Vinner (1981) describe a person's concept image for a particular concept as "the total cognitive structure that is associated with the concept." A number of studies delineate students' concept images of particular mathematical ideas (e.g., Artigue, 1992; Rasmussen, 2001; Wilson, 1993; Zandieh, 2000). In addition to work that uses concept image as its framing, we find useful studies that (whether they refer to it by the term concept image or not) detail student concept images of mathematical constructs using the construct of a *conceptual metaphor* (e.g., Lakoff & Nunez, 2000; Oehrtman, 2009; Zandieh & Knapp, 2006). This follows from the earlier work in cognitive linguistics of Max Black (1977), Lakoff and Johnson (1980) and Lakoff (1987). Following from this work, our assessment is that a person's concept image of a particular mathematical idea will likely contain a number of metaphors as well as other structures. Zandieh and Knapp (2006) provide an example of this for the concept of derivative.

In order to examine metaphors, we rely on metaphorical expressions. Lakoff and Johnson (1980) explain that, "Since metaphorical expressions in our language are tied to metaphorical concepts … we can use metaphorical linguistic expressions to study the nature of metaphorical concepts and to gain an understanding of the metaphorical nature of our activities (p. 456)." Our work will describe clusters of metaphorical expressions that allow us to highlight the connections or discrepancies between student conception of function and student conception of linear transformation.

Methods

The data for this report comes from interviews with 10 students who were just completing an undergraduate linear algebra course. The interviews were videotaped and transcribed and student written work was collected. The focus of the interview was to obtain information about students' concept image of function and their concept image of linear transformation and to see in what ways students saw these as the same or different. To this end we not only asked the students how they thought of a function or linear transformation, but also questions about characteristics that would be relevant to both functions and linear transformations such as one-to-one, onto, and invertibility. Several sample interview questions are provided below:

1. In the context of high school algebra, explain in your own words what a function is.
2. In the context of linear algebra, explain in your own words what a transformation is.
3. Please indicate, on a scale from 1-5, to what extent you agree with the following statement: "A linear transformation is a type of function."
4. In the context of high school algebra, give an example of a function that is 1-1 and one that is not 1-1. Explain.
5. In the context of linear algebra, give an example of a linear transformation that is 1-1 and one that is not 1-1. Explain.
6. Please indicate, on a scale from 1-5, to what extent you agree with the following statement: "1-1 means the same thing in the context of functions and the context of linear transformations."

We initially used grounded theory (Strauss & Corbin, 1994) to analyze student responses. As we refined our coding we noticed that the responses seemed to fall into three main types – *properties*, *computations*, and various *clusters of metaphorical expressions*. The details of these categories will be illustrated in the Results section. Coding with these categories followed an iterative cycle of coding by individual researchers, coming to consensus as to coding across individual researchers, and revising or refining the coding scheme as needed to more accurately reflect what we were seeing in the data. The next section documents the results of these deliberations.

Results

The main result of this paper comes in the form of a categorization of how students think about function and linear transformation. In order to compare students' concept images of function and linear transformation, we determined three main categories of tools students use to reason about these mathematical concepts: *properties*, *computations*, and *clusters of metaphorical expressions*. In this section we will provide examples of students reasoning with properties, computations, and each of the clusters. We will then provide sample results of how this categorization can be used to reveal important distinctions or connections between student conceptions of function and linear transformation.

Properties

While reasoning with the interview tasks, many students referenced a property of a function or linear transformation or a property of a feature associated with either such as a graph or a matrix. The property category refers to student statements that do not delve into the inner workings of the function or transformation. In the first example below, Andrew describes a function using a property about equations, and was coded P(equations). In the second example, Dana reasons about why a linear transformation is one-to-one by referring to linear independence P(li), presumably the fact that the columns of the associated matrix were linearly independent.

Adam: A function is an <u>equation with a variable</u>.

Donna: I said that was one-to-one because it's <u>linear independent</u>.

Computations

Students often drew upon computational language while reasoning through the interview tasks. We differentiated between computations that were done to carry out the function or transformation (labeled as C1), i.e., to get from the starting entity to the ending entity, and side computations done involving the function or transformation (labeled as C2), for example to compute the inverse function. In the first example, Ryan uses computational language (multiplication) to discuss how a linear transformation acts, which is indicative of C1. The second example shows Dana describing how to find the inverse of a 2x2 matrix. Her language (switch, make negative) is procedural and algorithmic, and *involves* the linear transformation but does not describe *how* the linear transformation acts.

Randall: A transformation is a <u>multiplication</u> of matrices that leads to a new image produced from the original matrix or vectors in the matrix.

Donna: Oh, I think you <u>switch</u> these two [points to entries on the off diagonal] and then probably <u>make this negative</u> [points to entries on the diagonal]. <u>Switch</u> those negatives.

Clusters of Metaphorical Expressions

We identified five different clusters of metaphorical expressions that students called upon when reasoning about function or linear transformation: *input/ output, traveling, morphing, mapping,* and *machine.* These five clusters share the common structure of a beginning entity, an ending entity, and a description about how these two are connected (see Fig. 1). Note that not all three parts of a structure must be stated by a student for the statement to be classified as part of a particular cluster.

Input/ Output

Input/ output involves an input, which goes into something, and an output, which comes out. This can be viewed from the point of view of the person ‘putting in’ the input and ‘taking out’ the output, and/or from the point of view of the function or transformation ‘accepting,’ ‘receiving’ or ‘taking’ an input and ‘returning’ or ‘giving’ an output. The first example shows Jordan using both of these perspectives in the same sentence. In the second example, George’s expression is from the point of view of the function.

Jerry: A function f of x = y means that <u>putting</u> x <u>inside</u> would <u>give</u> you a specific output, y.

Gabe: ... a function is an equation that <u>accepts</u> an input and <u>returns</u> an output based on that input.

Traveling

Traveling involves a beginning location being sent or moving to an ending location. Some phrases that we found to be indicative of this cluster were the use of ‘gets sent’, ‘goes to’, ‘moving, ‘reach’, ‘go back’, and ‘get to.’ These expressions were used almost exclusively when reasoning about linear transformation. We saw these expressions used to describe a pointwise change in location as well as a global move. In the first example, Andrew describes a transformation as a pointwise change in location, and in the second example George describes how transformations act more globally.

Adam: A transformation is <u>moving</u> a point or object in a <u>certain direction</u>.

Gabe: When you're in transformations, you'll always be able to <u>get back</u>. If a matrix is invertible, you should be able to <u>go both ways</u>.

Morphing

Morphing involves a beginning state of an entity that changes or is morphed into an ending state of the same entity. There must be a clear sense that the beginning entity did not simply move to the new location (ending entity), nor was it replaced by the new output (ending entity), but that there was actually a metamorphosis of the beginning entity into the ending entity. Morphing may be used pointwise by imagining one object changing, or globally by imagining a collection of objects changing. We found the phrases ‘become’, ‘transform’, and ‘change’ to be indicative of this cluster. In the following example, Dana uses morphing to explain what a transformation does to individual ‘things’.

Donna: Linear transformations to me are more or less something that <u>changes</u> something from one thing to another.

Mapping

Mapping involves a beginning entity, an ending entity, and a relationship or correspondence between the two. This cluster is most closely related to the Dirichlet-Bourbaki definition of function, and was not commonly used by students. We found the phrases ‘map’, ‘rule’, and ‘correspondence’ to be indicative of this cluster, as well as ‘per’ and ‘for’, as in there is one input for/per every output. This cluster was more commonly used in connection to function, but was used in relation to linear transformation as well. The following utterance is one of these uses:

Lawson: [A linear transformation is] a <u>rule</u> that assigns a given input to a certain output or image of the input.

Machine

Expressions in the machine cluster include a beginning entity or state, an ending entity or state, and a reference to a tool, machine or device *that causes* the entity to change from the beginning entity/state into the ending entity/state. A necessary component to expressions in this cluster is language that indicates that the function or transformation is performing the action on the entity. We found the phrases 'acts on' and 'produces' to be indicative of this cluster. In the first example, Noah indicates that the function is performing an action, and in the second example George focuses on the action of a linear transformation.

Nigel: *A function is an <u>operation</u> on something.*

Gabe: *Pretty much anything you toss in here, this is still that transformation should be able to <u>act on</u> it.*

Combined expressions

There are several general things to note when comparing across the clusters of metaphorical expressions for function and linear transformation. Each of the clusters has the same general structure and they are often used in combination in student reasoning. In particular since input/output focuses more on the beginning and ending entity, it can most easily be combined with each of the other clusters. However, students often flow from one cluster to another even in the same sentence. Below Brian combines the machine, input/output and morphing, while Landon combines the mapping, machine and input/output.

Brad: *I just remember when I was in middle school or elementary school or whatever, learning about functions, and learning about them as a <u>machine</u>, <u>you put something in</u>, and it <u>transforms</u> it to something else.*

Lawson: *Because it essentially does the same thing. So it's like, how I have here a <u>rule</u> that <u>assigns</u>, essentially a function is the same thing, you <u>put in an input</u>, and it <u>manipulates</u> that input and <u>turns it into an output</u>.*

Comment in relationship to the process-object dichotomy

The mapping cluster is closest to Sfard's (1992) or Breidenbach et al's (1992) object conception of function. The other clusters provide interesting nuances to our understanding of the process view of function.

Discussion: Using the Clusters to Analyze Student Understanding

In the Results section we provided details of the categories that came out of our analysis. Here we discuss some further results that illustrate the usefulness of a categorization of this type. The first two questions of the interviews directly addressed students' concept images of function and linear transformation (see questions 1 and 2 in the Methods section).

By comparing each student's responses to these questions, we can see that certain clusters of metaphorical expressions are called upon more frequently than others when reasoning about function or linear transformations (see Table 2). These results provide an interesting resource in understanding how students see function and linear transformation as similar or different mathematical concepts.

When discussing function, the input/output cluster (7 students) and the property of being an equation (4 students) were the most prevalent. By contrast when answering the same question for linear transformation, the morphing cluster (5 students) and the machine cluster (3 students) were most common. The traveling cluster (2 students) was only used by students answering this question for linear transformations. Notice also that all but one of the students used expressions from different clusters to answer this question for function than they did for linear transformation. However, when asked to indicate, on a scale from 1-5, to what extent you agree with the following statement: "A linear transformation is a type of function," all ten students marked 4 or 5 to indicate their agreement with that statement. Thus, these students may believe that function and linear transformation are related, but the initial evoked image for each are rather different for most students.

The third column in Table 2 shows how students responded when asked to elaborate on their answer of agreement with the statement. As students worked to reconcile their images, five students used additional clusters that they had not mentioned in response to the first two questions. Lawson provides a nice example of this. His answers to the three questions were as follows.

Lawson(function): A method that takes an <u>input</u> and <u>spits out an output</u>.

Lawson(linear transformation): A <u>rule that assign</u>s a given <u>input</u> to a certain <u>output</u> or image of the input.

Lawson(compare): I agree ... Because it essentially does the same thing. So it's like, how I have here a <u>rule that assigns</u>, essentially a function is the same thing, you <u>put in an input</u>, and <u>it manipulates</u> that input and <u>turns it into an output</u>. And that's essentially what a transformation I would say is, because <u>it transforms</u> something into something else.

In the first statement, Lawson's expression is from the input/output cluster, using the colorful language that the output is "spit out". In the second statement, Lawson again uses phrases from the input/output cluster but layers on the notion of "a rule that assigns" which is from the mapping cluster. In the third statement, Lawson states that "a rule that assigns" is the "same thing" as an input/output process, as well as this being "essentially" a transformation process (morphing cluster). He also brings out the idea that the function "manipulates" and a transformation "transforms" treating these entities as machines that act on something. Lawson seems comfortable using expressions from four of the metaphorical clusters in the same paragraph, speaking as if they are essentially synonymous, from his point of view.

Directions for future research

We find the analysis in terms of clusters of metaphorical expressions helpful for delineating important aspects of a student's concept image of function, linear transformation and how these are related. Further research will allow us to illuminate under what conditions each of the clusters are most likely to be used by students. In addition, we are interested in exploring how students interrelate the clusters. Because of the parallel structure of the five clusters -- each having a beginning entity, a middle, and an ending entity -- it is possible for students to layer the expressions on top of each other, in the way that Lawson does, blending aspects of one cluster with another, even though the metaphorical expressions refer to quite different images.

References

Artigue, M. (1992). Cognitive difficulties and teaching practices. In G. Harel, & E. Dubinsky (Eds.), The *concept of function: Aspects of epistemology and pedagogy* (pp. 109-132). Washington, DC: The Mathematical Association of America.

Black, M. (1977). More about metaphor. *Dialectica*, *31*, 433–457.
Breidenbach, D., Dubinsky, E., Hawkes, J. & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics, 23,* 247-285.
Carlson, D. (1993). Teaching Linear Algebra: Must the Fog Always Roll In? *The College Mathematics Journal, 24(1)*, 29-40.
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, 33* (5), 352-378.
Dorier, J.-L., Robert, A., Robinet, J., & Rogalski, M. (2000). The obstacles of formalism in linear algebra. In J.-L. Dorier (Ed.), On the teaching of linear algebra (pp. 85-124). Dordrecht: Kluwer.
Dreyfus, T., Hillel, J., & Sierpinska, A. (August, 1998). Cabri-based linear algebra: Transformations. Paper presented at the First Conference on European Research in Mathematics Education, Osnabrück, Germany.
Dubinsky, E. & Harel, G. (1992). The nature of the process conception of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy.* MAA notes, (pp. 85-106). Washington, DC: Mathematical Association of America.
Dubinsky, E., & McDonald, M. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton, M. Artigue, U. Krichgraber, J. Hillel, M. Niss, & A. Schoenfeld (Eds.), The teaching and learning of mathematics at university level: An ICMI Study (pp. 273-280). Dordrecht: Kluwer.
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,* 139-148.
Hillel, J. (2000). Modes of description and the problem of representation in linear algebra. In J.-L. Dorier (Ed.), On the teaching of linear algebra (pp. 191-207). Dordrecht: Kluwer Academic Publisher.
Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.
Lakoff, G., & Nunez, R. (1997). The metaphorical structure of mathematics: Sketching out cognitive foundations for a mind-based mathematics. In L. English (Ed.), *Mathematical Reasoning: Analogies, Metaphors, and Images* (pp. 21–92). Mahwah, NJ: Lawrence Erlbaum Associates.
Lakoff, G., & Nunez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.
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, Vol. 25 pp. 175-193). Washington, DC: Mathematical Association of America.
Oehrtman, M. (2009). Collapsing dimensions, physical limitation, and other student metaphors for limit concepts. *Journal for Research in Mathematics Education, 40*, 396-426.
Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' understandings of function. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 150-171). Washington, DC: Mathematical Association of America.

Portnoy, N., Grundmeier, T.A., & Graham, K.J. (2006). Students' understanding of mathematical objects in the context of transformational geometry: Implications for constructing and understanding proofs. *Journal of Mathematical Behavior 25*, 196-207.
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), 1-36.
Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification – The case of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (MAA Notes, no. 25, pp. 59-84). Washington, DC: MAA.
Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), On the teaching of linear algebra (pp. 209-246). Dordrecht: Kluwer Academic Publisher.
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.
Thompson, P. W. (1995). Students, functions, and the undergraduate curriculum. In *Research in Collegiate Mathematics Education,* Vol I. (pp. 21-44) Washington, DC: Conference Board on the Mathematical Sciences.
Vinner, S., & Dreyfus, T. (1989). Images and definition for the concept of function. *Journal for Research in Mathematics Education, 20*, 356-366.
Wilson, M. R. (1993). *One preservice teacher's evolving understanding of mathematical functions*. Paper presented at the Annual Meeting of the American Educational Research Association. Atlanta, GA.
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).
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.

Tables

Table 1: Structure of metaphorical expressions			
Cluster	Entity 1	Middle	Entity 2
Input/Output (IO)	Input(s)	Entity 1 goes/is put into something and Entity 2 comes/is gotten out.	Output(s)
Traveling (Tr)	Beginning Location(s)	Entity 1 is in a location and moves into a (new) location where it is called Entity 2.	Ending Location(s)
Morphing (Mor)	Beginning State of the Entity(ies)	Entity 1 changes into Entity 2.	Ending State of the Entity(ies)
Mapping (Map)	First Entity	Entity 1 and Entity 2 are connected or described as being connected by a mapping (a description of which First entities are connected to which Second entities).	Second Entity
Machine (Mach)	Entity(ies) to be processed	Machine, tool, device acts on Entity 1 to get Entity 2.	Entity after being processed

Table 2: How students initially explained function and linear transformation			
Student	**Function**	**Linear Transformation**	**How do you see these as the same?**
Adam	$P_{equation}$	Tr	IO, Mor, Tr, Comp
Brad	IO, Mor	IO, Mor	IO, Mor
Donna	IO	Mor	Mor, IO
Gabe	$P_{equation}$, IO	$P_{equation}$, Tr	IO, Mach, Mor
Jerry	IO	Mach	IO, Mach
Josh	Comp	Mor, Mach	Comp
Lawson	IO	Map, IO	Map, IO, Mor, Mach
Nigel	Mach	Mor	Mor, Mach
Nila	$P_{equation}$ IO	Mach	$P_{equation}$, P_{VLT}, Mach, Mor
Randall	$P_{equation}$, IO, Map	Comp, Mor	P_{VLT}