PROCEEDINGS OF THE 16TH ANNUAL CONFERENCE ON RESEARCH IN UNDERGRADUATE MATHEMATICS EDUCATION

EDITORS
STACY BROWN
GULDEN KARAKOK
KYEONG HAH ROH
MICHAEL OEHRTMAN

DENVER, COLORADO
FEBRUARY 21 – FEBRUARY 23, 2013

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

CITATION: In (Eds.) S. Brown, G. Karakok, K. H. Roh, and M. Oehrtman, *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education*, 2013, Denver, Colorado.

FOREWARD

As part of its on-going activities to foster research in undergraduate mathematics education and the dissemination of such research, the Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education (SIGMAA on RUME) held its sixteenth annual Conference on Research in Undergraduate Mathematics Education in Denver, Colorado from February 21 - 23, 2013.

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. Patrick Thompson, Dr. Koene Gravemeijer, Dr. Loretta Jones, and Dr. Keith Weber and the presentation of over 117 contributed, preliminary, and theoretical research reports and posters. In addition to these activities, faculty, students and artists contributed to displays on Art and Undergraduate Mathematics Education.

The Proceedings of the 16th 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 or unique insights into existing research programs.

Volume 2, *RUME Conference Reports*, includes the Poster Abstracts and the Contributed, Preliminary and Theoretical Research Reports that were presented at the conference and that underwent a rigorous review by at least three reviewers prior to the conference. Contributed Research Reports discuss completed research studies on undergraduate mathematics education and address findings from these studies, contemporary theoretical perspectives, and research paradigms. Preliminary Research Reports discuss ongoing and exploratory research studies of undergraduate mathematics education. Theoretical Research Reports describe new theoretical perspectives and frameworks for research on undergraduate mathematics education.

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

Sincerely,
Stacy Brown, RUME Conference Chairperson
Gulden Karakok, RUME Conference Local Organizer
Kyeong Hah Roh, RUME Program Chair
Michael Oehrtman, RUME Coordinator

PROGRAM COMMITTEE

Program Committee Chair	
Stacy Brown	California State Polytechnic University, Pomona
Local Organizer	
Gulden Karakok	University of Northern Colorado
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
Kyeong Hah Roh	Arizona State University
Elise Lockwood	University of Wisconsin at Madison
Michael Oehrtman	University of Northern Colorado
Vicki Sealey	West Virginia University
Hortensia Soto-Johnson	University of Northern Colorado
Craig Swinyard	University of Portland
Allison Toney	University of North Carolina–Wilmington
Joseph Wagner	Xavier University
Megan Wawro	Virginia Tech
Aaron Weinberg	Ithaca College

WITH MUCH APPRECIATION WE THANK THE FOLLOWING REVIEWERS

Belanp, Jason (PC)
Brown, Stacy (PC)
Engelke, Nicole (PC)
Karakok, Gulden (PC)
Keene, Karen (PC)
Lockwood, Elise (PC)
Oehrtman, Michael (PC)
Roh, Kyeong Hah (PC)
Sealey, Vicki (PC)
Soto Johnson, Tensia (PC)
Swinyard, Craig (PC)
Toney, Alison (PC)
Wagner, Joseph (PC)
Weinberg, Aaron (PC)
Wawro, Megan (PC)
Fukawa-Connelly, Timothy (PC)
Adiredja, Aditya
Austin, Jathan
Bagley, Spencer
Beisiegel, Mary
Borja-Tecuatl, Ileana
Boyce, Steven
Bubp, Kelly
Byrne, Martha
Cadwalladerolsker, Todd
Champion, Joe
Champney, Danielle
Chesler, Joshua
Code, Warren
Cook, John Paul
Cook, Samuel
Dawkins, Paul
Davis, Owen
Dibbs, Rebecca
Dolor, Jason
Dorko, Allison
Ellis, Jessica
Eubank, Alyssa
Fisher, Brian
Galle, Gillian
Glassmeyer, David
Grundmeier, Todd
Gula, Taras
Gupta, Shweta
Halani, Aviva
Hall, William
Hamdan, May
Hauk, Shandy
Horvath, Aladar
Jayakody, Gaya
Johnson, Estrella
Judd, April Brown
Kelton, Molly
Kim, Hyung
Kwon, Oh Hoon
Lai, Yvonne
Larsen, Sean
Lee, Miscun
Lew, Kristen
Lozano, Guadalupe
Martin, Jason
Meija-Ramos, Juan Pablo
Miller, David
Mills, Melissa
Moore, Kevin
Noblet, Kristin
Park, Jungeun
Paterson, Judy
Piatek-Jimenez, Katrina
Pilgrim, Mary
Powers, Robert
Rasmussen, Chris
Reinholz, Daniel
Rice, Lisa
Rizzo, Sarah
Samkoff, Aaron
Savic, Milos
Schmitz, Rebecca
Speer, Natasha
Stewart, Sepideh
Strand, Steve
Trigueros, Maria
Toth, Krista
Troupe, Jonathan
Wasserman, Nicholas
Weber, Eric
Weber, Keith
Wescoatt, Benjamin
Whitacre, Ian
Wijeratne, Chanakya
Wu, Linda
Yannotta, Mark
Zandieh, Michelle
Zazkis, Dov
Zazkis, Rina

CONFERENCE PARTICIPANT LIST

Aditya Adiredja
Valeria Aguirre Holguín
Dean Allison
Stephenie Anderson Dyben
Aina Appova
Inbar Aricha-metzer
Jathan Austin
Spencer Bagley
Rabindra Bajracharya
Greg Baker
Hyman Bass
Jason Belnap
Tetyana Berezovski
Sarah Bleiler
Tim Boester
Heather Bolles
Robert Borgersen
Ileana Borja-Tecuatl
Steven Boyce
Stacy Brown
April Brown Judd
Kelly M. Bubp
Ron Buckmire
Cameron Byerley
Martha Byrne
Dalia Challita
Joe Champion
Joshua Chesler
Kritika Chhetri
Warren M Christensen
Brian Arthur Christopher
Emily Cilli-Turner
Derron Coles
Noelle Conforti Preszler
John Paul Cook
Samuel Cook
Beth Cory
RaKissa Cribari
Jennifer Czocher
Owen Davis
Paul Dawkins
Rebecca-Anne Dibbs
Marki Dittman
JasonDolor
Allison Dorko
Tevian Dray
Irene M Duranczyk
Jess Ellis
Rob Ely
Nicole Engelke Infante
Brooke Ernest
Alyssa Eubank
Shawn Firouzian
Brian Fisher
Elizabeth Fleming
Bethany Fowler
Kristin Frank
Timothy Fukawa-Connelly
William Fulton
Michael Furuto
Gillian Galle
Shawn Garrity
David Glassmeyer
Jim Gleason
Erin Glover
Melissa Goss
Todd Grundmeier
Taras Gula
Shweta Gupta
Aviva Halani
Brent Hancock
Neil Hatfield
Shandy Hauk
Chuck Hayward
Joe Hernandez
Kendrice James
Wendy James
Gaya Jayakody
Garry Johns
Estrella Johnson
Elgin Johnston
Gulden Karakok
Brian Katz
Hyung Kim
Jeffrey J. King
David Kohler
Dave Kung
George Kuster
Oh Hoon Kwon
Yvonne Lai
Sean Larsen
Sandra Laursen
Younhee Lee
Kristen Lew
Stefanie Livers
Quin Wai Lo

Elise Lockwood
Tim Lucas
Michael Mack
Soofia Malik
Ami Mamolo
Corinne Manogue
Lisa Mantini
Jason Martin
Mike May, S.J.
John Mayer
William McCurdy
Pablo Mejía-Ramos
Kate Melhuish
Sandra Merchant
David Miller
Nathaniel Miller
Melissa Mills
Kevin C. Moore
Robert Moore
Stacy Musgrave
Reshmi Nair
Revathi Narasimhan
Kristin Noblet
Michael Oehrtman
Jerry Overmyer
Eric Pandiscio
Teo Paoletti
Jungeun Park
Thomas Parrish
Judy Paterson
Valerie Peterson
Costanza Piccolo
David Plaxco
Robert Powers
Jeff Rabin
Robert Raish
Daniel Reinholz
Dylan Retsek
Lisa Rice
Kitty Roach
Lee Roberson
Kim Rogers
Kyeong Hah Roh
Sarah Rozner Haley
Pooneh Sabouri
Milos Savic
Rebecca Schmitz
Vicki Sealey
Alee Seehausen
Annie Selden
John Selden
Ann Sitomer
John Smith
Hortensia Soto-Johnson
Natasha Speer
Dara Stepanek
Sepideh Stewart
Steve Strand
April Strom
Judy Sutor
Julie Sutton
Craig Swinyard
Jenna Tague
Matt Thomas
John Thompson
Pat Thompson
Michael Thoms
Dr. Anna Titova
Allison Toney
Krista Kay Toth
Marcy Towns
Jonathan Troup
Jen Tyne
Joseph Wagner
Quin Wai Lo
Aaron Wangberg
Nick Wasserman
Megan Wawro
Eric Weber
Aaron Weinberg
Thomas Edward Wells
Benjamin Wescoatt
Diana White
Nina White
Emilie Wiesner
Chanakya Wijeratne
Claire Wladis
Mark Yannotta
Hyunkyoung Yoon
Michelle Zandieh
Orit Zaslavsky
Dov Zazkis
Rina Zazkis
Ashley Zenisek
Bo Zhen

VOLUME 1

CONFERENCE PAPERS

TABLE OF CONTENTS

Volume 1: Conference Papers

BEST PAPER AWARD:

HONORABLE MENTION AWARD:

HONORABLE MENTION AWARD:

UTILIZING TYPES OF MATHEMATICAL ACTIVITIES TO FACILITATE CHARACTERIZING STUDENT UNDERSTANDING OF SPAN AND LINEAR INDEPENDENCE

Megan Wawro
Virginia Tech

David Plaxco
Virginia Tech

The purpose of this study is to characterize students' conceptions of span and linear (in)dependence and to utilize mathematical activities to provide insight into these conceptions. The data under consideration are portions of individual interviews with linear algebra students. Grounded analysis revealed a wide range of student conceptions of span and linear (in)dependence. The authors organized these conceptions into four categories: travel, geometric, vector algebraic, and matrix algebraic. To further illuminate participants' conceptions of span and linear (in)dependence, the authors developed a framework to classify the participants' engagement into five types of mathematical activity: defining, proving, relating, example generating, and problem solving. This framework proves useful in providing finer-grained analyses of students' conceptions and the potential value and/or limitations of such conceptions in certain contexts.

Key words: Span, Linear Independence, Linear Algebra, Mathematical Activity, Concept Image

The purpose of the study is to investigate student thinking about the important ideas of span and linear independence in linear algebra and to contribute to the body of knowledge regarding how individuals understand undergraduate mathematics. In particular, our research goals are:

1. To classify students' conceptions of span and linear (in)dependence.
2. To investigate how students use these conceptions to reason about relationships between span and linear (in)dependence.

The present study focused on interview data that elicited student reasoning about span and linear (in)dependence. We oriented our analysis through a grounded theory approach (Glaser & Strauss, 1967) in order to identify student conceptions of span and linear (in)dependence. We noticed that in coding students' conceptions, for which we made use of Tall and Vinner's (1981) construct of concept image, our analysis was facilitated by noting the type of mathematical activity in which the students were engaged as they were sharing their ways of reasoning. In other words, the interview question to which a student was responding had the potential of eliciting different aspects of the student's concept images. This is consistent with Vinner's (1991) notion of *evoked concept image*. For example, students' reasons why a claim was true or false revealed ways of thinking about the associated concepts differently than did their response to "how do you personally think about this concept?" As such, we identified within the data set five mathematical activities in which students engaged during the interviews: defining, proving, relating, example generating, and problem solving. We use these activities as a means through which to extend and further clarify the broader concept image framework. Thus, within this study we show how these mathematical activities can be used as a lens to further refine characterizations of students' concept images of span and linear (in)dependence.

Given this framework, our refined research objectives are (a) to investigate students' concept images of span, linear (in)dependence, and relationships between the two concepts and (b) to utilize the mathematical activities of defining, proving, relating, example generating, and problem solving, to provide insight into these concept images. Our results section details the four concept image categories that grew out of our data: travel, geometric, vector algebraic, and matrix algebraic. We also define the five mathematical activities and provide examples of how coordination of the frameworks informed analysis of student thinking.

Literature Review

There exists a growing body of research regarding multiple modes of description or reasoning within linear algebra. Two of the most cited are those by Hillel (2000) and Sierpinska (2000). Hillel suggested three possible modes of description for vectors and vector operations, namely geometric, algebraic, and abstract. The *abstract mode* utilizes language of generalized theory, including terms such as dimension, span, linear combination, and subspace. The *algebraic mode* uses concepts more particular to the vector space $\mathbb{R}^n$, such as matrix, rank, and systems of linear equations. Finally, the *geometric mode* uses language that is familiar from our lived experiences, such as point, line, plane, and geometric transformation (p. 192). Hillel details difficulties students have within a given mode (such as confusion potentially caused by describing vectors as both arrows and points, both a geometric description of vectors), as well as difficulties moving between modes (such as how the difficulty in change of basis problems within $\mathbb{R}^n$ may relate to switching between algebraic and abstract modes). Attributing them to the historical development of linear algebra, Sierpinska (2000) suggests three modes of thinking in linear algebra: synthetic-geometric, analytic-algebraic, and analytic-structural. The first mode focuses on spatial reasoning, the second on algebraic manipulation and representation, and the third on formal, theorem-based and axiomatic thinking.

In other research in linear algebra, Stewart (2009) and Stewart and Thomas (2010) coordinated APOS Theory (Dubinsky & McDonald, 2001) with Tall's (2004) Three Worlds of Mathematics to characterize the various possibilities for student understanding of linear independence, span, and basis according to the authors' genetic decomposition of the concepts. That is, the authors described possible conceptions along each level of the APOS framework—action, process, and object—and within the three worlds of mathematics—Embodied, Symbolic, and Formal. They classified and analyzed students' responses (from two different classes) using to this framework. The authors found that students in the more traditional course relied heavily on matrix manipulation (classified as a process-symbolic matrix conception at best) with little connection being drawn between the matrix manipulation and the associated concepts. Also, the participants in the study seemed to have limited understanding of the concept of linear combination, upon which the concepts of span and linear (in)dependence (and hence basis) rest. The authors recommended instruction that focuses on geometrically grounded development of the concepts (specifically linear combination), cautioned against too much reliance on the embodied aspects of linear algebra, and suggested that an appropriate balance would develop the concepts from a more geometric approach and relate these concepts at a more formal level.

This literature has informed and shaped the research we present in this paper. We drew inspiration with respect to the various modes of description that are possible within linear algebra as an aspect of student reasoning to be sensitive to within our analysis. While these studies expand our knowledge of student conceptions of span and linear (in)dependence, the current study differs in that our analysis of student conceptions are grounded with no a priori categorizations. As such, we also draw from the notion of concept image (Tall & Vinner, 1991), which has been utilized and adopted as an analytical framework to characterize students' conceptions of ideas within linear algebra. This approach allows categorizations to surface from the data through grounded theory, rather than through a priori classifications. For example, Zandieh, Ellis, and Rasmussen (2012) investigated students' conceptions of function and linear transformation and discovered five concept image categories for linear transformation within their data, such as morphing (an input morphs into an output) or machine (a transformation acts on an input to produce an output). Wawro, Sweeney, and Rabin (2011) document concept image categories for students' conceptions of subspace within linear algebra, namely part of whole, geometric object, and algebraic object. Although these categories were grounded in data, there exist points of compatibility with Hillel's and Sierpinska's groupings.

Setting and Participants

The data for this study come from a semester-long classroom teaching experiment (Cobb, 2000) conducted in an introductory linear algebra course at a large public university. Classroom instruction was guided by the instructional design theory of Realistic Mathematics Education (RME) (Freudenthal, 1991), with the goal of creating a linear algebra course that built on student concepts and reasoning as the starting point from which more complex and formal reasoning developed. The class engaged in various RME-inspired instructional sequences focused on developing a deep understanding of key concepts such as span and linear independence (Wawro, Rasmussen, Zandieh, Sweeney, & Larson, 2012), linear transformations (Wawro, Larson, Zandieh, & Rasmussen, 2012), Eigen theory, and change of basis.

The five students analyzed in this research - Abraham (a junior statistics major), Giovanni (a senior business major), Justin (a sophomore mathematics major), Aziz (a junior chemical physics major), and Kaemon (a senior computer engineering major) - participated in semi-structured individual interviews (Bernard, 1988) the week after final exams. Each interview lasted approximately ninety minutes. The purpose of the interview was to investigate how students reasoned about the concept statements that comprise the Invertible Matrix Theorem; the entirety of the interview protocol can be found in Wawro (2011). The current study considered only a portion of this data: students' conceptions of span, linear (in)dependence, and how they relate to each other. The interview questions analyzed in this study are given in Figure 1. Video recordings and transcripts of the interviews served as primary data sources, with all written work serving a secondary role.

Methods

Videos and transcriptions of the participants' responses to Questions 1a and 1b were iteratively analyzed. In the first analysis, the researchers focused on the logical progression of the participants' argumentation and what mathematical objects the participants attributed as 'acting' in different parts of their discussion (i.e., "the matrix spans $\mathbb{R}^3$" or "the vector moves in this direction"). A summarizing process that described the participants' general progression followed this analysis. A second analysis parsed out students' conceptions of linear (in)dependence and span, and we oriented our analysis through a grounded theory approach (Glaser & Strauss, 1967). We also separated general discussion of a concept from instances in which the interviewer directly asked the student to define the given concept. Quotes were drawn from the transcript and grouped by which concept the student was arguing with or describing. It was in this iteration of analysis that distinctions between types of activity became clear and led to the formation of the five categories of activity. In the next iteration of analysis, the researchers categorized student quotes according to these five activities and separated the quotes according to span, linear independence, or linear dependence. These quote collections were then compared for categorical similarities and differences. At every stage of this process, the two researchers continually questioned and challenged each other's decisions, such as motivation for choice of categorization or interpretation of a student's quote (Denzin, 1978).

"Suppose you have a 3 x 3 matrix *A*, and you know that the columns of *A* span $\mathbb{R}^3$. Decide if the following statements are true or false, and explain your answer:"

Question 1a

The column vectors of *A* are linearly <u>de</u>pendent.

Follow-ups. Skip if redundant:

- "How do you think about span?"
- "How do you think about what it means for vectors to be linearly <u>de</u>pendent?
- "How does linear <u>de</u>pendence relate to span of a set of vectors?

Question 1b

The row-reduced echelon form of *A* has three pivots.

Follow-ups. Skip if redundant:

- "How do you think about what a pivot is?
- "How do pivots of a matrix relate to span of a set of vectors?"

Figure 1. Interview questions analyzed for this study.

Results

The participants in this study used a variety of language to describe their understanding of span, linear (in)dependence, and how the two concepts relate to each other. We organized this variety into four concept image categories: travel, geometric, vector algebraic, and matrix algebraic (see Table 1). We also identified five mathematical activities in which students engaged during the interview: defining, proving, relating, example generating, and problem solving (Table 2). Within this section, we first detail the concept image category framework and the mathematical activity framework. We then illustrate the dual coding with an example from Justin's interview. The remainder of the section illustrates how the dual framework lends insight and nuance to characterizations of students' ways of reasoning about span and linear (in)dependence.

Table 1. Summary of the concept image framework for span and linear (in)dependence.

Category	**Description**
Travel	• Language indicative of purposeful movement • Captures notions of "getting to" or "moving to" locations in the vector space
Geometric	• Language indicative of spatial reasoning or graphical representations without use of travel-oriented language • Included sketches of vectors and/or discussion of objects such as lines and planes
Vector Algebraic	• Participants use operations on algebraic representations of vectors to describe concept • Includes linear combination of vectors written as *n* x 1 matrices or designated by variables (i.e., $2v + 3w$)
Matrix Algebraic	• Involves explicit attention to the form or properties of a matrix (e.g., size, actual values, pivots) • Participants focus on operations on matrices (e.g., Gaussian elimination)

Categories of student conceptions

The *travel* category captures students' description of span and linear (in)dependence in terms indicative of purposeful movement. While this category is consistent with spatial and geometric reasoning, it is more specific in that it captures notions of “getting to” or “moving to” locations in the vector space under consideration. The participants’ travel conceptions of span were indicated by phrases such as “everywhere you can get” (stated by Justin when describing the span of a set of vectors) or “the vectors can take you anywhere [in $\mathbb{R}^3$]” (stated by Giovanni when describing what it means for vectors to span $\mathbb{R}^3$). With respect to linear independence, participants’ travel conceptions included phrases such as “[the vectors] only move farther away” (Justin). A travel conception of linear dependence was generally indicated by phrases such as “then that would make that linearly dependent because I can, I can kind of get there and take that vector back” (Abraham), and “you can move one way on one vector, second way, and then take the third one back to the origin.” (Aziz). These were often given as the inverse of phrases used to describe sets of linearly independent vectors.

The *geometric* category is used to capture language indicative of spatial reasoning or graphical representations without use of travel-oriented language. This includes student explanations in which vectors are represented graphically on 2- or 3-dimensional axes, or when explanations include sketches or mention of objects such as lines, planes, or areas. For instance, when asked to discuss the span of three linearly dependent vectors he had given as an example, Kaemon stated, “So since it's, like you're on this one line, you can't really get all the combinations that are in these quadrants.” When asked how he thinks about what span means, Aziz replied, “Span is just the area that it covers. It could be a plane in $\mathbb{R}^3$, it could be a line in $\mathbb{R}^3$.” Most examples of the geometric category with respect to linear dependence consisted of students showing either two collinear vectors or three vectors placed head to tail to form a triangle with one vertex at the origin. For instance, Abraham constructed an example of three vectors in $\mathbb{R}^2$ and an associated sketch (see Figure 2) to explain geometrically why the three vectors are linearly dependent.

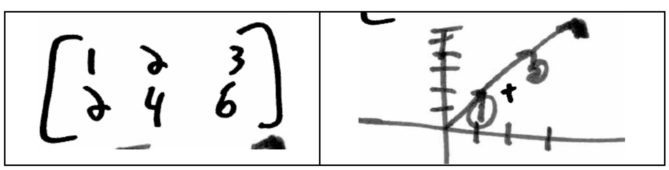


Figure 2. Abraham’s geometric explanation of three linearly dependent vectors.

The *vector algebraic* category captures participants’ use of operations on algebraic representations of vectors to describe span and linear (in)dependence. This includes scalar multiplication, vector addition, and linear combination of vectors written as n x 1 matrices, vectors designated by variables (i.e., $2\mathbf{v} + 3\mathbf{w}$), as well as the use of the equation $A\mathbf{x} = \mathbf{b}$. Vector algebraic conceptions of span included “every vector you can make with linear combinations of the columns” (Justin) and “in order to span $\mathbb{R}^3$, vectors have to be different” (Giovanni). Vector algebraic conceptions of linear independence consisted of some form of the notion that only the trivial linear combination of linearly independent vectors would equal the zero vector. An example of a student description of linear dependence that we coded as vector algebraic is, “But if it's linearly dependent, then there wouldn't be enough vectors because at least two of them are going to be maybe multiples of each other or just the zero vector” (Kaemon). Here, Kaemon attends to the quality of vectors being multiples of one another (as opposed to, say, vectors being collinear) to justify their linear dependence. One participant, Abraham, described linear independence as when the equation $A\mathbf{x} = \mathbf{b}$ has one unique solution. This notion is included in this category because Abraham tended to focus on the product as a linear combination of column vectors of A rather than on the matrix as an entity mapping $\mathbf{x}$ to $\mathbf{b}$ or as a system of linear equations (Larson & Zandieh, in press). Vector algebraic conceptions of linear dependence include when a nontrivial linear combination yields the zero vector and also the process of scaling one vector or taking a linear combination of vectors in order to produce a linearly dependent set.

Finally, the *matrix algebraic* category captures participants' explicit attention to the form or properties of a matrix, or on procedural operations on matrices. With respect to form or properties of matrices, instances in which students referenced the size or dimension of matrix when reasoning about span or linear (in)dependence were coded as matrix algebraic conceptions of span or linear (in)dependence. For instance, Giovanni's statement of, "I mean I would see it as being linearly dependent because you have more columns than rows" is an example of a matrix algebraic conception of linear dependence because he attended to the form of the matrix (more columns than rows). Instances in which participants made use of matrix-oriented algorithms such as Gaussian elimination through elementary row operations were also coded as matrix algebraic conceptions. For example, Aziz stated, "If it [a 3x3 matrix] doesn't reduce to the identity then it means it doesn't span all of $\mathbb{R}^3$." In fact, in our data, this was most prevalent occurrence of a matrix algebra conception—that is, participants' reliance on the row-reduced echelon form of a matrix either equaling the identity matrix or containing a row or column of zeros.

Types of Mathematical Activity

The construct of types of mathematical activity emerged from our grounded analysis of the data. As we analyzed students' understanding of span and linear (in)dependence in light of the concept image construct, we found ourselves continually drawn to notice the *type of activity* in which students were engaged as they responded to the interview questions. For instance, if a student spoke of span using a phrase such as "get everywhere," was that student engaged in explaining how span related to linear independence, explaining how he thought about the concept of span itself, or some other activity? As such, we identified five mathematical activities within the interview data: defining, proving, relating, example generating, and problem solving (see Table 2). We contend that considering the five mathematical activities provides insight into a student's understanding of a concept. We can consider these facets of a student's interaction with the world based on what s/he understands a concept to be. These activities do not always occur in isolation. Furthermore, an activity may arise naturally based on the interview prompt or may occur spontaneously. Here we provide descriptions and examples of each category of mathematical activity.

Table 2. Summary of the mathematical activity framework.

Mathematical Activity	**Definition**
Defining	The act of describing a concept's essential qualities
Proving	The act of providing a justification to a claim
Relating	The act of comparing, contrasting, or explaining relationships between different concepts or between different interpretations of the same concept
Example Generating	The act of creating cases of certain concepts or properties (e.g., a set of three linearly dependent vectors in $\mathbb{R}^3$)
Problem Solving	The act of engaging in some calculation or reasoning with a specific goal to determine a previously unknown result

We use the term *defining* to mean the act of describing a concept's essential qualities. During the interviews, students were not asked to create definitions for concepts that were new to them, but rather to explain their notion of a concept that had been defined during their linear algebra course. As such, this use of defining may be of a slightly different connotation than the discipline-specific practice of defining (e.g., Zandieh & Rasmussen, 2010). For example, Giovanni, when prompted to explain in general how he thought about span, replied, "The way that I think of span is just being able to reach anywhere in, like in $\mathbb{R}^3$, like in that dimension, like you're able to get to all points in $\mathbb{R}^3$." Also, if students spontaneously (i.e., without prompting) described a concept, we coded that as a "defining" activity. An example of this is given in the next section, in which Justin spontaneously describes the concepts of linear independence and linear dependence in service of explaining how the two concepts are similar and different.

We use the term *proving* to mean the act of providing a justification to a claim. This reasoning process may be of various levels of mathematical rigor, and it may be carried out for the participant's personal conviction or to convince the interviewer. This justification may be of a variety of forms, such as a chain of reasoning or a coordination of more than one justification to support a claim. As such, we use "proving" similarly to Harel and Sowder's (1998) use of "the process of proving," which included the subprocesses of ascertaining and persuading. For example, in response to Question 1a in Figure 1, Giovanni states:

> That's false, because in order to have a 3 by 3 matrix that spans all of $\mathbb{R}^3$, the column vectors have to be linearly independent. And so that's how, basically that's the definition, so I think of it that way.

In response to the same question, Aziz asserts, "The matrix A is a 3 by 3 … and since it spans all of $\mathbb{R}^3$, the columns, the column vectors of A are linearly independent." As implicitly illustrated by this quote, the proving code is not meant to determine levels of acceptability or correctness of the student's statement; rather, it is solely meant to convey when students are engaged in the mathematical activity of proving.

We use the term *relating* to denote any participant activity that compares, contrasts, or explains relationships between two concepts or between different interpretations of the same concept. The activities of proving and relating are similar, but distinct. We distinguish between these activities by attending to the overarching purpose of the student's utterance. For instance, participants might provide a statement of two concepts' relationship in support of claims during a proving activity. In the next section, we provide an example of relating activity. The activity of *example generating* denotes when participants create cases of certain concepts or properties (e.g., a set of three linearly dependent vectors in $\mathbb{R}^3$). As with the other activities, this may be prompted by the interviewer explicitly or spontaneously done by the interviewee. Finally, the activity of *problem solving* is engaging in some calculation or reasoning with a specific goal to determine a previously unknown result; this is consistent with the NCTM Process Standard definition of problem solving (NCTM, 2000). For example, to glean more information about a student's conception of span or linear (in)dependence, the interviewer occasionally would pose a problem about span or linear (in)dependence to a student and have him solve it. The activity of problem solving did not occur as frequently as the other codes within the data set; we attribute this to the nature of the interview questions, which were created to have students engage in justifying a true/false conclusion and explaining their understanding of the related concepts. Thus, by virtue of the interview questions, the mathematical activities of defining, proving, and relating were most common in the data set.

Coordinated Analysis: Using the Two Frameworks

We first illustrate, with one short section of transcript, how we made use of the two frameworks within our analysis. The given transcript comes from a portion of Justin's interview in which he was describing the differences and similarities between linear independence and linear dependence:

Justin: They're exactly the same as in it's where you can get, it's just different rules in how you get there, and if you can get back and whatnot.

Interviewer: Can you say some more about the difference in the rules?

Justin: The difference is huge! You know, uh, the difference is that with independence, you can only go farther away and maybe kind of sideways, but you can't come back. Dependence, you can make it back to where you started.
And, um, so even though I know they're not the same at all really, it's just they are the same because it's describing where you can get to, just it says different things about them.

In lines 1-2, we code Justin's discourse as the mathematical activity of *relating*, and the underlined portions indicate an image category of *travel.* In response to the interviewer's inquiry in line 3, Justin elaborates. In line 4, Justin *relates* linear independence and dependence by saying "the difference is huge." In lines 5-7, he then, in order to substantiate that claim, engages in the activity of *defining.* That is, he gives what, to him, is an essential quality of linear independence by stating, "You can only go farther away and maybe kind of sideways, but you can't come back." He then engages similarly for linear dependence, stating, "You can make it back to where you started." Both of these activities of defining are marked with language consistent with *travel* imagery, as indicated by the underlined portions of lines 5-7. Finally, we code Justin's discourse in lines 8-10 as *relating* as he concludes his explanation of the similarities and differences between linear independence and dependence, all which again using language consistent with the *travel* image category.

The remainder of the results section is dedicated to three examples that illustrate how coordinated analyses using the two frameworks lends insight into the research results. That is, we consider how three participants (Abraham, Kaemon, and Aziz) engaged in Question 1a (see Figure 1). Each participant's response provides a unique opportunity to observe how coordinating the two frameworks is useful. First, we show how Abraham's problem solving activity provides insight into how he related his varied conceptions of span. We then show how Kaemon's restricted conceptions of span and linear dependence possibly prevented him from meaningfully relating the two concepts. Finally, we describe Aziz's process of 'refiguring out' how his notions of span and linear dependence are related through an example generating activity.

Abraham. We begin by providing the first portion of transcript of Abraham's initial response to Question 1a:

Abraham: That would be false. Let's see. [7-sec pause] So if, the way I think of it is, if it, if it's spanning a 3 by 3 matrix [draws the outer brackets for a matrix], um, then it's going to have like, you know, three pivot positions here [writes three ones in on the diagonal of his matrix]. And for like a square matrix, I just think like if this is three pivots in each row, then it's also going to be, automatically going to be 3 pivots in each column. And that way you're always going to have a linearly independent set of...of, um, like x equals something, y equals something, z equals something, because of that. And then that’s, so that's going to be, basically a unique solution for every output. So let's see, so [writes $A\mathbf{x} = \mathbf{b}$], for every $\mathbf{b}$ there's a unique $\mathbf{x}$ vector.

After (correctly) commenting that the given statement is false, Abraham began by writing an empty matrix with 1’s on the main diagonal (lines 1-4) and claimed that “if it’s spanning a 3x3 matrix” (line 2), there would be three pivots down the diagonal. We note that Abraham did not directly address whether the matrix is given in this form or whether it must be manipulated to fit this form (via elementary row operations). He went on to claim that such a set of vectors would be linearly independent because the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution (lines 6-11). We categorize these initial statements as the mathematical activity of proving. We further coded this as matrix algebraic because his language indicates that Abraham utilized matrix algebraic conceptions of span and linear independence within his justification. Later, however, while explaining why he drew the three 1’s down the diagonal of the matrix, Abraham said, “if it's spanning something, it kind of needs to, I think of it like it needs to go in every direction.” Here, Abraham provided an essential quality of span, so this was coded as a defining activity. Notice also that the language of this defining activity indicates a travel conception of span, distinct from his previous use of a matrix algebraic conception of span.

The interviewer probed Abraham to further explain the connection between the 1’s and “going in every direction”:

Interviewer: How is it that those 3 pivot positions allow you to do those 3 directions?

Abraham: I go in this direction and then I can kind of pivot up, you know to go to this direction … so I can kind of go any direction, based on these 3 pivots, like pivoting in each direction, to get to any point.

Interviewer: That's interesting, I like that language. So like if, say you had an example like, I don’t know, 3,4,5, can you explain how you would use that language of pivoting to get to 3,4,5?

Abraham: I'm, I’m saying that I could use, you know, three [writes “3<1,0,0>”] of that one, a linear combination of this guy [writes “+ 4<0,1,0>”] and then this [writes “+ 5<0,0,1> =”]. I could write that out as a linear combination of 3,4,5 [writes “<3,4,5> to the right of the equals sign]. And so you can see, I can do that for a lot of different vectors, and so if I think of it in these terms, the basis vectors, that I could, you know, whatever I put here [points to the ‘3’ in his first vector’s scalar], is going to be at the top [points to the ‘3’ in the <3,4,5> vector]. So I can get to any point there, I can get to any point there [repeats the analogous gestures for the “4” and ‘5’] and so forth to get to that point.

Within lines 19-28, Abraham showed, for the given vector $\begin{bmatrix}3\\4\\5\end{bmatrix}$, how the columns in the matrix with spanning column vectors (standard basis vectors) could be used in linear combination “to get to that point.” In lines 23-28, Abraham alluded to being able to do this for any given vector. The mathematical activity within lines 23-28 is coded as problem solving because the vector $\begin{bmatrix}3\\4\\5\end{bmatrix}$ is a specific instantiation of a general vector in $\mathbb{R}^3$, and because Abraham engages in his demonstration through an algorithmic process. What is notable about this process is that Abraham’s actions focused on each column of the matrix as an individual vector, whereby he calculated a scalar multiple for each vector so that the resulting linear combination would yield the vector $\begin{bmatrix}3\\4\\5\end{bmatrix}$. That is, Abraham’s focus shifted from matrix algebraic and travel conceptions of span (in lines 1-11) to a vector algebraic conception of span, within which his actions were oriented around linear combinations of vectors (lines 19-28).

Through the problem solving activity, Abraham demonstrated relationships between his travel and matrix algebraic conceptions of span. This could be thought of as a “meta-relating” activity, within which – although the immediate purpose of his work was to solve a closed-ended problem – Abraham used a linear combination of vectors to relate how the matrix with spanning column vectors (his matrix algebraic conception of span) indeed reaches everywhere in $\mathbb{R}^3$ (his travel conception of span). We can view this as Abraham’s active coordination of his many conceptions of span, in which the various parts of his understanding come together to produce a meaningful (to Abraham) way for him to think about what it means for vectors to span a vector space. Thus, by using the dual framework to consider Abraham’s engagement in this specific mathematical problem, the researchers gained deeper insight into his conceptions of span and linear independence than when merely stating an essential quality – an insight into how *he* operationally coordinated his own varied understanding.

Kaemon. In his initial response to Question 1a, Kaemon said the following:

Kaemon: Ok. [6-sec pause] Uh, I say that this is false because the key is that it's 3 by 3 and it’s, it says you know that the columns of A span $\mathbb{R}^3$, so, um, like the minimal amount of vectors you could have to span $\mathbb{R}^3$ is at least three. But if it's linearly dependent, then there wouldn't be enough vectors, because at least two of them are going to be maybe multiples of each other or just the zero vector. So that's why I say it is false.

Kaemon engaged in this proving activity using a matrix algebraic conception of span (focusing on the size of the matrix, lines 1-3) and a vector algebraic conception of linear dependence (how individual vectors relate to each other, lines 4-6). The interviewer then asked Kaemon to generate an example of a matrix with linearly dependent column vectors, to which he responded with the matrix $\begin{bmatrix}1 & 2 & 0\\1 & 2 & 0\\1 & 2 & 0\end{bmatrix}$. This example generating activity shows Kaemon attended to the relationships between the individual vectors, also indicating a vector algebraic conception of linear dependence. When asked to explain why these columns vectors are linearly dependent, Kaemon stated that these vectors are “all on the same line” and added, “they won't be able to span, it's dependent.” This first statement shows that Kaemon is attending to a geometric conception of linear dependence when discussing the example he generated. The second statement is coded as relating activity, although there is no clear indication of what conception of span and linear dependence Kaemon is attending to when stating this relationship.

The interviewer then asked Kaemon to define span, prompting the following dialogue:

Kaemon: I think span is, I think it's all the possible linear combinations of the matrix. So since it's, like you're on this one line, you can't really get all the combinations that are in these quadrants [moves pen around the coordinate system], so that's how I think of it.

Interviewer: Ok. So you just said a second ago the way you think of span is all of the linear combinations of the columns? That's good. Um, how do you think about linear dependence in general?

Kaemon: Um, dependence for me is just, first I just try to look with, if there's a matrix to see if it's, like, if it's already reduced, then I see if there's a variable, or I see that they're multiples, that they're 0 vector, just something to show it's dependent. And then if I can't find that, like, right away, then maybe I'll try to then, I don't know, try to reduce it, whatever, just until I could figure it out.

We categorize lines 7-10 as defining activity indicative of a vector algebraic conception of span. This is a slight shift from Kaemon's earlier focus on the dimensions of the matrix (taken to mean height and width of an array, not necessarily the formal, vector algebraic *dimension*, lines 1-2). This dialogue also indicates a shift from vector algebraic and geometric conceptions of linear dependence to a matrix algebraic conception, because his focus changed from scalar multiples of vectors to the row-reduced echelon form of a matrix. Notice, though, that part of Kaemon's defining activity explicitly addresses scalar multiples of column vectors, as well as the presence of a zero vector, but necessarily requires the elementary row operations involved in row reduction before considering vector algebraic relationships between the column vectors. Kaemon similarly defined linear independence in the context of a row-reduced matrix. We note here that this definition of linear dependence (and hence linear independence) is restricted to a single algorithm carried out on a specific, given matrix.

In Kaemon's initial response (lines 1-6), he correctly answered that the statement in Question 1a was false, but he never explicitly stated or demonstrated a relationship between the concepts of span and linear independence. We notice that Kaemon had multiple conceptions of span and linear (in)dependence but failed (at least during the interview) to meaningfully coordinate these conceptions through any of his mathematical activity. We attribute part of this limited ability to relate the two concepts to Kaemon's matrix algebraic conceptions of linear independence and linear dependence. Specifically, Kaemon's definition of each concept necessitated a given matrix upon which he could carry out the elementary row operations. We note here that Kaemon alluded to, but never engaged in, problem solving activity, which may have provided deeper insight into how he might have coordinated his conceptions (as in Abraham's case). Also, consider Kaemon's conceptions of linear dependence: collinearity when discussed geometrically, scalar dependence when discussed vector algebraically, and inclusion of the zero vector in a row-reduced matrix when discussed matrix algebraically. These are restricted forms of the mathematical definition, for instance, neglecting coplanarity and its higher dimensional analogues or restricting the context to a problem-solving situation, dependent on the row reduction algorithm.

Aziz. In our discussion of Aziz, rather than concentrating on his response to Question 1a, we focus on a specific episode wherein he established a relationship between linear dependence and span. Aziz described a specific type of linear dependence in which "two, three completely different matrices (sic) that are not in combinations of each other, but they combine in a way to equal zero." He was contrasting this type of linear dependence with an earlier example he had generated that contained three vectors, two of which were collinear. Formally, Aziz is trying to generate a set of three coplanar vectors in $\mathbb{R}^3$, none of which is pair-wise collinear with another. Aziz then engaged in the following example generating activity:

Aziz: So *C* equals say 2,2,2; this is 1,2,3. And then, you know, I don't know how to calculate it right now.
Interviewer: Ok. What would you, just off the top of your head, what would you do—
Aziz: 2, 2$\mathbf{v}$, $2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = 0$.
Interviewer: Hmm. Ok I—
Aziz: So that way you can move one away on one vector, second way and then take the third one back to the origin.

Initially, Aziz attempted to generate specific column vectors in $\mathbb{R}^3$ but stalled (lines 1-2). He then attempted to describe what such a linear combination would be, but he was unable to use this algebraic representation with the first two vectors he had produced to generate the third vector. Throughout this example generating activity, Aziz's discussion focused on linear combinations of vectors, first representing them as columns in a 3x3 matrix, and then writing them as the abstract vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. This indicates a vector algebraic conception of linear dependence under two different representations. Aziz then used travel language to describe the linear combination (lines 6-7), indicating a travel conception of linear dependence. He followed up this activity with a geometric representation of this travel conception of span (Figure 3). We note that this is not explicitly a geometric conception of linear dependence, which would use geometric representations of vectors but would not use the travel language of movement.

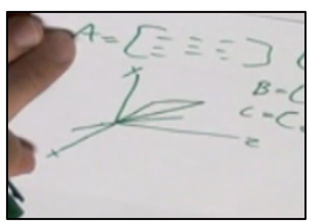


Figure 3. Aziz's geometric representation of three linearly dependent vectors.

When asked whether these vectors span $\mathbb{R}^3$, Aziz was perturbed (lines 8-10). Aziz's mathematical activity shifts here from example generating to relating activity (lines 14-15) and proving activity (13-14). He outlined a seeming contradiction, which we reorder and paraphrase for clarification: vectors that span $\mathbb{R}^3$ "move in three different directions" (9-10), a set of linearly dependent vectors does not span $\mathbb{R}^3$ (14-15), and this example shows vectors that "move in three different directions and get back to the origin [are linearly dependent]" (12-14). Notice that the second of these statements is possibly false if the set under consideration contains more vectors than the dimension of the vector space.

Aziz: They're linearly dependent. Um...that's a problem I always thought, because if it's they move in 3 different directions, they should technically span $\mathbb{R}^3$, but I never got clarification on that.

Interviewer: Say a little bit more about what's confusing to you about that?

Aziz: Because if they uh, they move different directions from each other, [Interviewer: Hm-mm] but they're linearly dependent, because you can use a combination of all 3 to get back to the origin. So linear dependence means it doesn't span all of $\mathbb{R}^n$. Right?

Aziz made sense of this seeming contradiction by pointing out that the linearly dependent vectors he generated move three different directions on the same plane and not outside of that plane. Drawing a distinction between two ways of "moving in three directions," Aziz was able to reason that his generated vectors did not span $\mathbb{R}^3$. In this episode, Aziz began with an example generating activity that utilized a vector algebraic conception of span. After shifting his example generating activity to draw on a travel conception of linear dependence, Aziz was perturbed by a seeming contradiction to a relationship he held between linear dependence and span. Finally, Aziz made sense of this contradiction through his proving activity that drew upon a travel conception of each.

Conclusion

This paper describes our work in categorizing students' concept images of span and linear (in)dependence and our use of the construct of mathematical activity to provide insight into these conceptions. We note that the concept image categories that arose may be an artifact of the type of instruction and curriculum that these students experienced; as such, we would expect that with different data sources (such as whole class discussion rather than individual interview data) or different participants (such as students from a more advanced, proof-based linear algebra course), additional or alternative categories for student conceptions would arise from the data. We also note that the five types of mathematical activity within our framework—defining, proving, relating, example generating, and problem solving—are not meant to be exhaustive; rather, these five activities were determined from analysis of this small data set. Analysis of classroom data or problem-solving interviews, for instance, would likely give rise to additional types of mathematical activity. As such, our future work involves a further examination and refinement of the framework of mathematical activity as a way to gain insight into students' conceptions of mathematical ideas. In addition, we also plan to examine additional data (e.g., classroom video data at the whole-class and small-group level, mid-semester interviews) of these same five students to gain a more complete analysis of their understanding.

This being said, we have found the coordination of the two frameworks to provide deeper, richer descriptions of student's conceptions of linear (in)dependence and span that would not have been possible through only one of the frameworks. Specifically, our results utilizing these frameworks have shown how students' engagement in different activities provides useful descriptions of the varied facets of the students' conceptions. For instance, Abraham's defining and proving activity, while mathematically appropriate and sound, did not provide insight into how he coordinated his varied conceptions of span. It was not until his problem solving activity that he demonstrated his deeper understanding of span – more precisely, until he was able to coordinate his matrix algebraic and travel conceptions of span. Similarly, Kaemon's defining activity demonstrated a restricted understanding of linear independence and linear dependence –

restricted in the sense that his understanding of the concepts depended on a specific context in which he could perform an algorithm to test for a specific trait in the row-reduced matrix. Without Kaemon engaging in such problem solving activity during the interview, we are unable to determine if his conceptions of linear independence and linear dependence are informed by more meaningful connections (as we saw with Abraham).

Generally, for each of our participants, a given participant utilized different conceptions when engaging in different activities. This is distinct from the notion of using different representations (e.g., vectors as arrows, *n*-tuples, or abstract **v**) in different activities. For instance, Aziz used *n*-tuple and abstract representations with his vector algebraic conception of linear dependence, but he used a geometric representation with his travel conception of linear dependence – all of which occurred during example generating activity. Further, it was not until Aziz represented linear dependence geometrically while utilizing his travel conceptions of span and linear dependence that he was able to discover and make sense of a seeming contradiction of his perceived understanding of the relationship between the two concepts. From this, we see that the activity framework extends the focus of research that examines students' conceptions into activities other than defining.

Given the power that a coordinated analysis has provided within this initial small data set, we have begun research that further explores this framework. In particular, we designed a new interview protocol for semi-structured individual interviews that is composed of tasks that (a) purposefully aim to engage students in the five various mathematical activities, and (b) make use of a variety of vector representations (e.g., arrows, specific vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$, or abstract vectors such as **v** and **w**). The mathematical content of the interview protocol focuses on student understanding of span, linear (in)dependence, and relationships between the two concepts. As such, Question 1a (analyzed in this paper) is included in the new interview protocol. We conducted interviews with seven, first-year, honors STEM majors, and the analysis of these data is ongoing.

References

Bernard R. H. (1988). *Research methods in cultural anthropology*. London: Sage.

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.

Denzin, N. K. (1978). *The research act: A theoretical introduction to research methods*. New York: Mcgraw-Hill.

Dubinsky, E., & McDonald, M. A. (2001). APOS: Constructivist theory of learning in undergraduate mathematics education research. *The Teaching and Learning of Mathematics at University Level, 7*(3), 275-282.

Freudenthal, H. (1991). *Revisiting mathematics education.* Dordrecht: Kluwer Academic Publishers.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine Publishing Company.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In E. Dubinsky, A. Schoenfeld and J. Kaput (eds.), *Research in Collegiate Mathematics Education*, III, 234-283. Providence, RI: American Mathematical Society.

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.
Larson, C. & Zandieh, M. (in press). Three interpretations of the matrix equation $A\mathbf{x} = \mathbf{b}$. *For the Learning of Mathematics*.
Lay, D. C. (2003). *Linear algebra and its applications* (3rd ed.). Reading, MA: Addison-Wesley.
National Council of Teachers of Mathematics. (2000*). Principles and standards for school mathematics.* Reston, VA: NCTM.
National Council of Teachers of Mathematics. (2000*). Principles and standards for school mathematics.* Reston, VA: NCTM.
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.
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.
Skemp, R. R. (2006). Relational understanding and instrumental understanding. *Mathematics teaching in the middle school*, 12(2), 88.
Stewart, S. (2008). Understanding linear algebra concepts through the embodied symbolic and formal worlds of mathematical thinking, Unpublished Doctoral Thesis, Auckland University.
Stewart, S. & Thomas, M. O. J. (2010). Student learning of basis, span and linear independence in linear algebra. *International Journal of Mathematical Education in Science and Technology*, *41*(2), 173-188.
Tall, D. O. (2004), The three worlds of mathematics. *For the Learning of Mathematics*, *23*(3), 29–33.
Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics, with special reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151–169.
Wawro, M. (2011). Individual and collective analyses of the genesis of student reasoning regarding the Invertible Matrix Theorem in linear algebra. *Dissertation Abstracts International, 72*(11). Retrieved May 14, 2012 from Dissertations & Theses: Full Text. (Publication No. AAT 3466728).
Wawro, M., Larson, C., Zandieh, M., & Rasmussen, C. (2012, February). *A hypothetical learning trajectory for conceptualizing matrices as linear transformations*. Paper presented at the 15th Conference on Research in Undergraduate Mathematics Education, Portland, OR.
Wawro, M., Rasmussen, C., Zandieh, M, Sweeney, G., & Larson, M. (2012). An inquiry-oriented approach to span and linear independence: The case of the Magic Carpet Ride sequence. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies. 22(*8), 577-599.
Wawro, M., Sweeney, G., & Rabin, J. M. (2011). Subspace in linear algebra: Investigating students' concept images and interactions with the formal definition*. Educational Studies in Mathematics, 78*, 1-19.
Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In D. Tall (Ed.), *Advanced mathematical thinking* (65-79)*.* Dordrecht, Netherlands: Kluwer.
Zandieh, M., Ellis, J. & Rasmussen, C. (2012, February). Student Concept Images of Function and Linear Transformation. In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education,* pp. 320-328. Portland, OR.
Zandieh, M., & Rasmussen, C. (2010). Defining as a mathematical activity: A framework for characterizing progress from informal to more formal ways of reasoning. *Journal of Mathematical Behavior*, *29*, 57-75.

MATHEMATICIANS' EXAMPLE-RELATED ACTIVITY WHEN PROVING CONJECTURES

Elise Lockwood, Amy B. Ellis, & Eric Knuth
University of Wisconsin – Madison

Examples play a critical role in mathematical practice, particularly in the exploration of conjectures and in the subsequent development of proofs. Although proof has been an object of extensive study, the role that examples play in the process of exploring and proving conjectures has not received the same attention. In this paper, results are presented from interviews conducted with six mathematicians. In these interviews, the mathematicians explored and attempted to prove several mathematical conjectures and also reflected on their use of examples in their own mathematical practice. Their responses served to refine a framework for example-related activity and shed light on the ways that examples arise in mathematicians' work. Illustrative excerpts from the interviews are shared, and four themes that emerged from the interviews are presented. Educational implications of the results are also discussed.

Keywords: Examples, Proof, Mathematicians

Introduction

Proof is a perennial topic in mathematics education, and while many cases have been made for its importance among students across age levels (e.g., Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; Knuth, 2002a, 2002b; Sowder & Harel, 1998; Yackel & Hanna, 2003; CCSS, 2010; NCTM, 2000), students' difficulties with proof seem to be persistent (Kloosterman & Lester, 2004). Some researchers have suggested that students' struggles with understanding the nature of evidence and justification may be due, in large part, to their views concerning the role and status of examples. In particular, students tend to be overly reliant on examples and often infer that a (universal) mathematical statement is true on the basis of checking a number of examples that satisfy the statement (e.g., Healy & Hoyles, 2000; Knuth, Choppin, & Bieda, 2009; Porteous, 1990). Much of the current literature on teaching proof in school mathematics underscores the goal of helping students understand the limits of such example-based reasoning (e.g., Harel & Sowder, 1998; Stylianides & Stylianides, 2009; Zaslavsky, Nickerson, Stylianides, Kidron, & Winicki, 2012) and typically characterizes example-based reasoning strategies as obstacles to overcome. However, given the essential role examples play in mathematicians' exploration of conjectures and subsequent proof attempts, we propose that example-based reasoning strategies should not be positioned only as barriers to negotiate. Indeed, the field may benefit from a greater understanding of the ways in which those who are adept at proof, such as mathematicians, critically analyze and leverage examples in order to support their proof-related thinking and activity. While the role of examples in learning mathematics more generally has received attention in the literature (e.g., Bills & Watson, 2008), there is still much to learn about the specific roles examples play in exploring and proving conjectures. In this paper, we examine mathematicians' example-related activity as they explore and develop proofs of several conjectures. We report themes that arose during the interviews and discuss potential implications for the teaching and learning of proof. The research question we attempt to answer in through this study is: What roles do examples play for expert mathematicians in the context of exploring and proving mathematical conjectures?

Relevant Literature

As a preliminary note, we briefly provide definitions for what we mean by *examples* and *proof*. In this paper, we define an example as Bills and Watson (2008) do, as "any mathematics object from which it is expected to generalize" (p. 78). In defining *proof*, we draw on Harel and Sowder's (1998) definition, which is "the process employed by an individual to remove or create doubts about the truth of an observation" (p. 241). Harel and Sowder further distinguish between two kinds of activity associated with proving – *ascertaining,* which they define as "the process an individual employs to remove her or his own doubts about the truth of an observation" (p. 241), and *persuading,* which is "the process an individual employs to remove others' doubts about the truth of an observation"(p. 241). We follow their lead and consider proof as considering both kinds of activities. For further clarification, we use *inductive* and *empirical* interchangeably to refer to example-based arguments. We also use *deductive* and *proof* interchangeably to refer to arguments comprised of a series of logically connected assertions that one makes to justify a mathematical claim.

It is generally accepted that students' understandings of mathematical justification are "likely to proceed from inductive toward deductive and toward greater generality" (Simon & Blume, 1996, p. 9); that is, students' justifications are expected to progress from empirical arguments to proofs. However, in this progression, caution must be made so that students do not view examples as constituting a proof or as being a valid substitution for a proof. There is a trend in the literature, then, toward helping students understand the limitations of examples as a means of justification and thus recognize the need for a proof (e.g., Sowder & Harel, 1998; Stylianides & Stylianides, 2009; Zaslavsky, Nickerson, Stylianides, Kidron, & Winicki, in press). Perhaps because of this predominant view, less work has been done that investigates useful ways in which students can actually use examples productively as they reason about conjectures and explore proof.

In recent years, there has been an increase in attention on examples in mathematics education literature. This is evidenced in part by a special issue on the topic of examples in *Educational Studies in Mathematics* (Bills & Watson, 2008). In their introduction to this issue, the editors note the goals of the issue as being to "raise the profile of this field as an important domain of research; focus attention on some issues concerning the role and effective use of examples in teaching and learning mathematics; bring these into a coherent articulation from which future directions for research may be formulated" (p. 78). In this issue and elsewhere, researchers on examples have included the study of: example generation in the context of novel definitions (Zazkis & Leikin, 2008), example spaces (Goldenberg & Mason, 2008), non-examples (Tsamir, Tirosh, & Levenson, 2008), examples in teaching (Rowland, 2008), and the importance of prior knowledge when comparing examples (Rittle-Johnson, Star, & Durkin, 2009), and more. This growing interest in examples in mathematics education research suggests that focusing on examples is a potentially fruitful line of research, worthy of study in its own right.

It is clear that examples play a critical role not only in mathematicians' development of and exploration of conjectures, but also in their subsequent development of proofs of those conjectures. Epstein and Levy (1995) contend that "Most mathematicians spend a lot of time thinking about and analyzing particular examples," and they go on to note that "It is probably the case that most significant advances in mathematics have arisen from experimentation with examples" (p. 6). Several mathematics education researchers have accordingly examined various aspects of the interplay between example-based reasoning activities and deductive reasoning activities among both mathematicians and mathematics students (e.g., Antonini, 2006; Buchbinder & Zaslavsky, 2009; Iannone, Inglis, Mejia-Ramos, Simpson, & Weber, 2011; Harel, 2008; Knuth, Choppin, & Bieda, 2009; Weber, 2008, 2010). More specifically, Harel (2008) notes that, "the empirical proof scheme does have value. Examples and non-examples can help to generate ideas or give insight" (p. 7). He goes on to caution, "The problem arises in contexts in which a deductive proof is expected, and yet all that is necessary or desirable in the eyes of a student is a verification by one or more examples" (p. 7), but the point is that Harel acknowledges some potential value in example-related activity. Additionally, in a discussion of proofs by mathematical induction, Harel (2002) suggests that certain process pattern generalizations, even if based in specific examples, can suggest transformational and not just empirical proof schemes. Weber (2008) has similarly reported that mathematicians at times drew on example-based arguments in determining the validity of an argument. Proponents of the notion of cognitive unity (e.g., Garuti, Boero & Lemut, 1998; Pedemonte, 2007) also distinguish between argumentation (conjecturing activities) and proof, and make the case that argumentation could involve work with examples and may help in the proving process. In addition, some researchers have begun to notice that perhaps students' are more sophisticated in their uses of examples than teachers and researchers give them credit for. For example, undergraduate students studied by Weber (2010) seemed to realize that even though they, at times, used examples when proving, they recognized that examples were not sufficient as proofs. At the middle school level, Knuth and his colleagues (e.g., Ellis, et al., 2012) have begun to explore students' uses of examples when examining conjectures, recognizing that perhaps students display some sophisticated uses of examples that contribute to meaningful mathematical reasoning, even reasoning that may contribute directly to proof.

Two additional example-related studies grounded the present study. First, Antonini (2006) interviewed advanced graduate students, asking them to generate examples with specific mathematical properties. His work yielded an initial categorization for producing these strategies (trial and error, transformation, and analysis). While our study ultimately differs in its emphasis, his categorization of these mathematicians' strategies provides a starting point for developing our themes. Indeed, Antonini had pointed out that

> "Further research is also needed to study how the identified strategies are intertwined with processes enacted in different situations where subjects produce examples. We also believe that these strategies may be useful to observe processes of production of examples in tasks involving a careful exploration in order to produce conjectures and proofs" (p. 63).

Iannone, Inglis, Mejia-Ramos, Simpson, and Weber (2011) studied the effects of example generation on proof production in undergraduate students. They build upon Antonini's (2006) framework discussed above and categorized students' strategies in example generation. Iannone et al. indicate that they were surprised to find that example generation did not seem to have a positive effect on proof production tasks. Their work underscores the complexity of studying examples, and they suggest that there is a need for further research in the area of examples (specifically to better understand example generation).

In the work presented in this paper, we build upon the studies mentioned above with the belief that there may be more to the role of examples in proof than simply signifying an unsophisticated line of reasoning. These aforementioned studies suggest that by conducing research that specifically targets how examples might be used meaningfully in the proving process, researchers may gain a more robust understanding of the role of examples in proof. To accomplish this, we carefully examine mathematicians, who we take to be experienced provers. Our work extends such research by focusing particularly on the role of examples as expert mathematicians' explore and attempt to prove mathematical conjectures.

Theoretical Framework

Theoretically, the study presented in this paper builds directly upon a framework developed by Lockwood, et al. (2012) that categorizes types of examples, uses of examples, and example-related strategies reported by mathematicians in a large-scale open-ended survey. The framework is presented in Figures 1, 2, and 3 below. This framework guided the coding of the interviews in this study and served to situate the themes presented below. To some extent, the interviews were conducted in order to test the viability of these surveys, providing some qualitative support of the quantitative findings. We thus were interested in the extent to which the interview results aligned with this existing framework of mathematicians' example-related activity.

Example Type	Definition
Simplicity	Expert appeals to an easy, simple or basic example. Includes "trivial" and "small."
Counterexample /Conjecture Breaking	Expert picks an example that might disprove the conjecture. The expert might explicitly say "a counterexample," but this can also be inferred.
Complex	Expert picks a complex example in order to test whether the conjecture holds for tricky ones; synonyms include "non-nice," "non-trivial," or "interesting."
Easy to Compute	Expert chooses an example that is easy to manipulate. The difference between this code and "Simple" is that the expert says something about working the example out.
Properties	Expert takes into account some specific mathematical property – he or she might reference a "property" or "features," or might mention particular properties.
General/Generic	Expert uses general or generic examples, or describes examples that are seen as representative of a general class of cases or otherwise lack special properties.
Boundary Case	Expert picks an extreme example or number, or a "special" case, such as the identity.
Familiar/Known case	Expert chooses an example with which he or she is familiar, or in which properties related to the conjecture are already known.
Unusual Examples	Expert picks an unusual number, which would be described as something that does not come up often. "Rare," "obscure," "strange," and "weird" are also synonyms.
Random	Expert describes the example as randomly chosen; this includes mathematical randomness, such examples chosen with a random number generator.
Exhaustive	Expert looks for "all" of the examples in an exhaustive manner. This can be by testing all possible examples or by using a computer.
Common	Expert describes the example as typical, common, or one many would choose.
Dissimilar Set	Expert indicates that he or she purposely selects a variety of types of examples.

Figure 1 – Types of Examples

Example Use	Definition
Check	Expert selects examples to make a judgment about the correctness of a conjecture; "test," "verify," and "check" are all synonyms.
Break the Conjecture	Expert tries examples to break the conjecture; this can include specifically looking for a counterexample.
Make Sense of the Situation	Expert uses an example to deepen his or her understanding of why the conjecture might be true or false, or to gain mathematical insight.
Proof Insight	Expert indicates that his or her production of examples (or counterexamples) might have a direct bearing on understanding how to prove the conjecture.
Generalize	Expert mentions using the example to generalize or to allow the expert to work in a more general situation.
Understand Statement of the Conjecture	Expert uses an example to better understand the statement of the conjecture.

Figure 2 – Uses of Examples

Strategy (Frequency)	Subcategory (Frequency)	Definition
Multi-Stage Example Exploration	*Increasing in Complexity*	Expert begins with simple or easy examples and shifts to more complex or complicated examples.
	Increasing in Extremity	Expert begins with simple or typical examples and builds to boundary cases, special cases, or conjecture-breaking cases.
	Increasing in Generality	Expert begins with simple or special examples and shifts to more general or generic examples.
Property Analysis	*Supporting or Non-Supporting*	Expert attempts to determine the properties of examples that either support the conjecture in particular ways, or to determine the properties of examples that do not support the conjecture.
	Test Cases	Expert analyzes special test cases depending on the critical properties of the examples related to the conjecture.
Analysis of Related Proof Activities	*Generalize From Attempts to Prove*	Expert proves related conjectures or lemmas and attempts to generalize to the conjecture at hand. Or, proves the conjecture for specific examples and attempts a general proof. Or, attempts to generalize from examples related to unsuccessful proof attempts.
	Generalize From Attempts to Disprove	Expert attempts to generalize from examples that disprove related conjectures. Or, attempts to generalize from properties of examples that failed to disprove the conjecture.
Systematic Variation	*Known Case Adjustment*	Expert takes a known case and makes small adjustments to the example's properties, inputs, or characteristics.
	Multiple Property Variation	Expert varies multiple properties or characteristics simultaneously or independently in a systematic fashion.

Figure 3 – Example-Related Strategies

While the literature and framework presented above provides theoretical backing for our paper, we make a final comment about the perspectives that have influenced our work. We agree with Weber (2008) who notes that, "Investigations into the practices of professional mathematicians should have a strong influence on what is taught in mathematics classrooms (e.g., RAND, 2003); the link between the behaviors of mathematicians and the teaching of mathematics, however, is not straightforward" (p. 451). We are working under an assumption that we, as mathematics education researchers, might intrinsically care about what expert mathematicians do. This does not mean that we necessarily believe that all novices should automatically adopt expert practices (this could be problematic for a number of reasons), nor that mathematicians' activity is somehow inherently superior to what a novice might do. However, it stands to reason that better understanding how mathematicians think about and use examples can better inform us about how examples can potentially be utilized. Particularly given the fact that there is more to learn about how examples can be effectively used in proof, looking to mathematicians seems to be a reasonable starting point. Additionally, there are a number of other instances in which researchers have examined expert mathematicians in a variety of contexts including as problem solving (e.g., Carlson and Bloom, 2005) and proof (e.g., Savic, 2012).

Methods

The data presented in this paper come from interviews that were conducted with six mathematicians as they explored and attempted to prove several mathematical conjectures (Figure 4). Five of the mathematicians have a doctorate in mathematics, and one has a doctorate in mathematics education; all are currently faculty in university mathematics departments. All of the mathematicians were given Conjectures 1 and 2, and three each did one of Conjectures 3 and 4, which were randomly assigned. After working on each conjecture, the mathematicians were asked clarifying questions about their work. In addition, at the end of the interview they were asked reflective questions about their example-related activity, both that they had done during the interview, and more generally in their personal work. They were given approximately 15-20 minutes to explore each conjecture; although typically they were not able to complete proofs for each of the conjectures in the time allotted, they were able to make progress toward that end. (Note that our interest was in their example-related activity while exploring and attempting to develop proofs, not in the proofs they may have produced given more time.)

Conjecture 1
Let S be a finite set of integers, each greater than 1. Suppose that for each integer n there is some $s \in S$ such that $\gcd(s,n)=1$ or $\gcd(s,n)=s$. Prove that there exist $s,t \in S$ such that $\gcd(s,t)$ is prime.

Conjecture 2
Let n be an even positive integer. Write the numbers 1, 2, …, n^2 in the squares of an n x n grid so that the kth row, from left to right, is $(k-1)n+1,(k-1)n+2,\dots,(k-1)n+n$. Color the squares of the grid so that half of the squares in each row and in each column are red and the other half are black.
Prove or disprove: For each coloring, the sum of the numbers on the red squares is equal to the sum of the numbers on the black squares.

Conjecture 3
Let S denote the set of rational numbers different from {-1, 0, 1}. Define $f:S\to S$ by $f(x)=x-1/x$.
Prove or disprove: $\bigcap_{n=1}^{\infty} f^{(n)}(S)=\varnothing$, where $f^{(n)}$ denotes f composed with itself n times.

Conjecture 4
All the numbers below should be assumed to be positive integers.
Definition. An abundant number is an integer n whose divisors add up to more than $2n$.
Definition. A perfect number is an integer n whose divisors add up to exactly $2n$.
Definition. A deficient number is an integer n whose divisors add up to less than $2n$.

Conjecture 4a. A number is abundant if and only if it is a multiple of 6.
Conjecture 4b. If n is deficient, then every divisor of n is deficient.

Figure 4 – The conjectures given to the mathematicians

Conjectures 1-3 were taken from Putnam Exams, and Conjecture 4 was adapted from tasks in Alcock & Inglis (2008). We chose these problems for two primary reasons. First, the conjectures were accessible to the mathematicians (regardless of their area of expertise), but were not so clearly obvious that they could be proven immediately. Second, the conjectures were also accessible to the interviewer, allowing her to follow the mathematicians' work as well as to ask meaningful follow-up questions. While the choice of such conjectures may result in something of an inauthentic situation for the mathematicians (in that these conjectures are not representative of their personal research), the choice did enable us to observe what mathematicians do as they actually explore and attempt to prove conjectures.

The interviews were transcribed, and a member of the research team analyzed them using the aforementioned framework (Lockwood et al., 2012). The process involved coding both mathematicians' observable example-related activity and their reflections. The entire research group reviewed data excerpts that were difficult to code. These codes served to refine the initial framework, and the organizing of the codes in turn resulted in a number of themes about mathematicians' example-related activity in exploring and proving conjectures (Strauss & Corbin, 1998). We present these themes as the major results of this paper, as they shed light on how people who are adept at proof interact with examples as they consider conjectures.

Results

In this section, we share four main themes that arose from our analysis of the interview data. Some of these themes (Themes 1 and 2) confirm what is seen in the literature (e.g., Schoenfeld, 1985; Weber, 2008), providing further details that flesh out prior findings. Theme 3 also has some precedence in the literature (e.g., Weber, 2008), but our particular example provides additional insight into the precise nature of how a proof might emerge out of a particular example for an expert mathematician. We highlight Theme 4 highlight it both as a particular aspect of the relationship between examples and proof, but also as a metacognitive action that suggests mathematicians' overall awareness of their proving activity. These themes not only illuminate the role examples play in the proving process for mathematicians, but also suggest implications regarding the role examples might play in classroom settings.

Theme 1 – Domain expertise and knowledge of mathematical properties directly influence example choice

The notion of domain knowledge affecting mathematical practice is brought up by a number of researchers. For example, in his discussion of problem solving resources, Schoenfeld (1985) notes "the successful implementation of heuristic strategies in any particular domain often depends heavily on the possession of specific subject matter knowledge" (p. 92). Within proof, Weber (2008) found that "some participants' validation standards and strategies were dependent on their familiarity with the domain that they were investigating" (p. 447). Regarding examples, Rittle-Johnson et al. (2009) indicate that prior knowledge was an important for middle school students' as they compared examples. Also, a number of other researchers (e.g., Sandefur, et al., 2013; Stylianides & Stylianides, 2009) argue for the importance of familiarity and domain in examples via Watson & Mason's (2005) notion of students' *personal example space*. The below insights gleaned from our interviews corroborate such findings that report the importance of domain knowledge, but we also provide some new ways of thinking about how prior mathematical knowledge affects example choice in the proving process.

We see this theme affect example choice in two primary ways, first in terms of domain knowledge (or lack thereof), and second regarding knowledge of mathematical properties. In terms of domain knowledge, four of the mathematicians noted that context and familiarity have a direct impact on their selection of examples, often enabling them to make well-informed choices. Specifically, in some cases mathematicians indicated that if they were working in domain they knew well, they would regularly draw upon familiar, or "stock," examples. For instance, on the deficient number problem (Conjectures 4a and 4b), Dr. Hickson clearly used his familiarity with the fact that 6 is a perfect number to make progress on that task, as seen in Dr. Hickson's exchange below.

Dr. Hickson: Conjecture 4a: A number is abundant if and only if it is a multiple of six. Hmmm ok. So an example immediately comes to mind. Six is a perfect number and so that's going to be false if you are allowed to take a trivial multiple of six. So….

Interviewer: …Ok. And that you knew six was a perfect number from experience.

Dr. Hickson: Yeah, that one I just happen to know.

In other cases, if the domain was less familiar, the mathematician might rely on examples to make sense of the conjecture. This is exemplified in Dr. Aldridge's reflection. Here he indicates that in a familiar domain he might simply launch into a proof without having to consider examples, but that when he is "completely clueless" he tries to generate examples. This exchange provides evidence that mathematicians may use examples to ground their work in something concrete, particularly when they need to make sense of a given situation or conjecture. Dr. Aldridge's activity described below puts a slight twist on the manipulation stage of the *manipulating-getting-a-sense-of-articulating* (e.g., Sandefur, et al., 2013), because while he talks about manipulating a concrete example to get a sense of the conjecture, he does so in a domain with which he is less familiar (although the example he may choose may be familiar).

Interviewer: … Can you describe the role of examples in your work with mathematical, mathematical conjectures? How do you choose then? Do you have strategies for example-related activity? Like if you were to have to reflect on how you would use examples?

Dr. Aldridge: So, well, first of all, it depends on the, the domain. I mean, there's some domains when I know, very familiar with all of the, like the more algebraic, formal techniques…and I can kind of recognize if it's a situation where I can actually get by without even really understanding…the problem, because I can just throw the tools at it…and it'll fall out… Other than that, I usually try to, I go in a couple different ways. Especially if I'm completely clueless about what's going on, then I will usually use an example to try to figure out what's the conjecture is saying.

Another feature of the role of examples was that the mathematicians capitalized on their understanding of mathematical properties as they selected their examples. This supports related findings by Weber (2008), who reports that some mathematicians drew upon mathematical properties as they validated proofs (p. 441). In our interviews, the mathematicians took into account the domain to which the conjecture pertained (such as number theory or algebra), and they used that knowledge to pinpoint examples with certain properties. Their mathematical expertise came through as they spoke about mathematical features of their examples, such as choosing a number that is highly divisible or creating a set with no primes. This emphasis on properties came out most frequently with Conjecture 1, as the mathematicians tried to consider examples or counterexamples of the conjecture. In this case, the mathematicians clearly drew upon their knowledge of mathematical topics such as primeness, common divisors, the fundamental theorem of arithmetic, etc.

As an example, Dr. Leonard constructed a set {4, 8, 12, 20} in an attempt to derive a counterexample. He had recognized that a counterexample must not have primes in it, and the excerpt below highlights his consideration of specific mathematical properties as he attempted to construct a possible counterexample and proceed with the problem. In considering what might be needed to make a counterexample, Dr. Leonard displays knowledge of elementary number theory as he carefully selects four numbers that are not prime and that all have a composite number as a greatest common divisor. Facility with specific mathematical properties enabled him to make sophisticated decisions in constructing an example.

Dr. Leonard: The greatest common divisor between the two of them [looking at the statement of the conclusion] is not prime…Okay, it would have to be some set like 4 [writes {4, 8, 12, 20}]. That would be…their greatest common divisor is not prime.

In another example of the use of mathematical properties, again on Conjecture 1, we highlight Dr. Hickson's language, which is interspersed with words like *prime, relatively prime, divides*, etc. These are not especially sophisticated mathematical ideas, but he clearly has easy access to and facility with these concepts, and they are having an effect on how he is constructing his examples in this case.

Dr. Hickson: Let's see if a number n were relatively prime to let's say a power of two and a power of three….I was going to say I just need there to exist a number in S that the number is relatively prime to. So suppose the number has no 2s in it, then it could be relatively prime to a 4. If it had no 3s in it would be relatively prime to a 9. And if it's not relatively prime to either of those then it's got at least one 2 in it and at least one 3 in it. And if it fails to be relatively prime to both of those numbers, then it would actually have to have 2 and a 3 in it. So, you could say 6 divides it. So that may be an example of a set that contains no primes and yet satisfies the conditions.

In both of the data excerpts above, we see that mathematical properties are a driving aspect that influences how mathematicians chose and constructed examples. It is worth speculating whether or not this differs from school-aged students' work with examples. It may be the case that with limited knowledge of mathematical properties (particularly compared with mathematicians), students simply may not have the mathematical knowledge to develop examples and make well-informed decisions about examples as the mathematicians.

Theme 2 – Multiple examples can lead to meaningful patterns, resulting in conjecture generation and proof development

This theme builds on work by Harel (2002), who discussed patterns in his presentation of DNR-based instruction of mathematical induction, and Weber (2008), who reported a handful of instances in which mathematicians accepted a statement as valid based on a pattern that emerged from several examples. In our study, five of the mathematicians demonstrated an explicit awareness of the relationship between examples and patterns in their work. As they worked through the conjectures, they suggested two different ways in which multiple examples could yield meaningful patterns that could shed light on their work with a conjecture.

First, Dr. Leonard reflects a use of patterns that has more to do with the act of conjecturing. He notes that rather than using examples while trying to prove a conjecture (as he did in the interviews), in his own research patterns of examples tend actually to lead to the formulation of conjectures.

Dr. Leonard: First of all, I wouldn't come up with a conjecture without an example…Again, some conjectures can come from just putting in a number of examples. You start thinking maybe it is true because I've seen it enough times, and I cannot conceive clearly of an argument of why this fails.

Dr. Hickson also shared some insightful comments about a certain way in which generation of examples and patterns can emerge in his own work. Below, he describes a building up of patterns that serve the purpose of shedding some light on an existence proof. He indicates that examples can serve a very specific purpose for him in this way, and that patterns he notices in multiple examples can give some particular insight about the nature of the family of objects he needs to construct (and prove that exists).

Dr. Hickson: So, this might be of interest to you because it's a quirky use of examples. So I believe that a certain thing exists for all values of n. And um the existence proof, it baffles me. But as far as I look, these things exist…I need to come up with an infinite family of them. And so it's very worth it to me to look for examples because I'm hoping by looking at the examples I'll notice a pattern of them…And I'll be able to say, 'oh actually you can build these things in this way and it will always satisfy the conditions.' But I feel like without the inspiration coming from the examples I won't know what that is going to look like.

This theme is also tangentially related to the problem solving heuristic of solving smaller, similar problems (e.g., Lockwood, 2013; Polya, 1945, Schoenfeld, 1979), which suggests a building up of instances of a given phenomenon or situation. The common idea between the theme and this heuristic is that someone might use multiple examples (or smaller instances of a problem or situation) to formulate a pattern in a systematic way.

Theme 3 – Examples can lead to proof insights, both into whether the conjecture is true or false, and into how a proof might be developed

There were two ways in which mathematicians seemed to use examples to gain some insight into their proving process. First, examples served to inform whether or not a given conjecture might be true or false, and at some point each mathematician used an example to decide whether he should go about trying to prove or disprove the conjecture. As noted above, Harel and Sowder (1998) refer to this as *ascertaining,* which they define as "the process an individual employs to remove her or his own doubts about the truth of an observation" (p. 241). Below, Dr. Wells' exploration with multiple examples suggests to him that the conjecture seems true.

Interviewer: And, at this point, do you have a sense of whether you think it's true?

Dr. Wells: It seems pretty true. I've constructed two simple examples. Uh, and I can see, it looks like I see a pattern here because, you know, the squares are just off by one, and so that's why you're getting the sums are equal.

Another episodes with Dr. Wells highlights this phenomenon as well, as he concludes that coming up with a particular example that satisfied the conjecture suggested that he should go about proving (as opposed to disproving) the conjecture.

Interviewer: Okay. Okay, so, um, came up with a small example. And when you came up with that example were you trying to make sense of the conjecture? Were you looking for whether it might be true or false?

Dr. Wells: Well, first I was trying to make sense of the hypothesis, to make sure the hypothesis was even possible.

Interviewer: Okay. Okay. Right.

Dr. Wells: And then once I convinced myself that they hypothesis was possible, in the example of the set being two… Then I was able to use that example to, um, to show that in the case of that example, the conjecture worked.

Interviewer: Okay. Okay.

Dr. Wells: And so that convinced me to try to prove the conjecture rather than to disprove it.

Second, examples served a richer purpose than simply shedding light on whether a statement was true or false. On several occasions mathematicians used specific features of an example in order to make significant steps toward a proof. In these instances, the mathematicians seemed to ground their thinking in a particular example, and by manipulating that example they developed an idea for how a more general proof might develop. A number of researchers have suggested that proofs can emerge from work with examples (e.g., Sandefur, et al., 2013; Weber, 2008), and more specifically in terms of a generic example (e.g., Pedemonte & Buchbinder, 2012). We contribute further support to these findings, and we also add to it by showing a mathematician using actually using an example to formulate a proof in the context of exploring a conjecture. We see in this example that Dr. Aldridge's work with a particular example allowed him to make general arguments and ultimately formulate a valid proof of the conjecture.

As a particularly rich example of how examples led to proof insight in the interviews, we highlight Dr. Aldridge's work Conjecture 4b as he tried to prove the contrapositive of the statement (that a number with not deficient factors must itself not be deficient). Dr. Aldridge had been examining what he called "test cases," in which he drew upon the perfectness of 6 to examine numbers in which 6 was a factor. His rationale for this is seen below.

Dr. Aldridge: And then the real reason why I went after it with examples, not so much that I thought these would be counterexamples, as I thought they would be good test cases. And they'd maybe give me a feel for how, more information as to maybe why this is true.

Interviewer: Okay, and what do you mean by test case?

Dr. Aldridge: Um, test case because the six, like I said before is perfect. So it's going to be, it's a, it's a pretty decent, uh, example of maybe, it's, so if anything has a chance to be a divisor that's not deficient inside of number that is deficient… I would guess it would be a perfect number.

Continuing to focus on 6, after trying to see if 6*2 and 6*3 would have to be abundant, he chose an example of 6*11. While working through this example, he had the following insight.

Dr. Aldridge: It's almost like you get, like a duplication of the perfect-ness of six that shows up in this piece here.

Interviewer: Okay, how so?

Dr. Aldridge: So, so, like this one, two, three adds up to six. Eleven, twenty-two, thirty-three actually adds up to sixty-six. So I'm feeling like I probably ought to be able to prove that this is a true statement.

His work with this example not only confirmed that he thought he could prove the conjecture, but using 6*11 as a generic example ultimately led him to a correct sketch of a proof. For the sake of space we briefly summarize his proof, but in his work he wrote out the sketch of proof and simultaneously referring back specifically to his example in doing so. He was able to prove the contrapositive, arguing that that if a factor of *b*, *a*, is not deficient, then it has factors *f1* through *fk* whose sum is greater than *a.* Then *d*f1* through *d*fk* must also be factors of *b* that are distinct from those factors of *a* (and since *f1* through *fk* are strictly less than *a, d*f1* through *d*fk* must be distinct from *b*). He noted that the sum of *d*f1* through *d*fk* must be greater than or equal to *d*a*, which is itself already a copy of *b*. The sum of *b*'s factors, then, includes *d*f1* through *d*fk,* which is greater or equal to *d*a*, and *b* itself. This is greater than or equal to two copies of *b*, and thus by the definition of deficient, *b* itself cannot be deficient. What is most interesting to us is not that he proved the conjecture, but rather the precise role that his example 6*11=66 played in his development of this proof. In reflecting on his proof, he made several statements that highlighted the importance of the example. Specifically, the nature of the multiplication by 11 allowed him to see that certain factors (the multiples of 6) would show up in the complete list of factors. While this is a property that he asserted, "is clearly always going to work out," he acknowledged that the nature of the number 11 made that particularly salient for him. The structure of the example enabled him to recognize a key piece of the proof.

Theme 4 – There is a back and forth interaction between proving and disproving

A final theme is that mathematicians seemed to report a complementary relationship between proving and disproving. All six interviewees discussed the role of counterexamples in their proving process, noting that as they attempt to develop a proof, they engage in a back and forth process of formulating a proof and considering counterexamples. They described starting out by attempting to prove a conjecture, but then may get stuck, stop, and search for a counterexample. This search for (or inability to find) a counterexample might then provide insight into the development of the proof. An example of this is seen in Dr. Barton's reflection about his work with examples. He articulates in great detail exactly how the search for a proof and a counterexample interact in his work on a conjecture.

Dr. Barton: You're trying to prove something and you go ahead and you try to prove it. And you realize that you're stuck at some point…Here's this gap. I start saying let's try, out of that gap, to build a counter example…Then you spend some time trying to build that object. And if you can't, then you try to sort of distill why can't you? And do the reasons why you can't build that, does that now fill in the gap in your proof? If it does, great. You've now pushed your proof further or maybe you've completed the proof entirely. And if it doesn't, then it refines what...the counter-example would have to look like…And so it's this sort of back and forth trying to use that. You know build a counter example and the failure or success of that to go back and look at what that says about your proof. And that dynamic back and forth can sometimes bear some fruit.

In reflecting on the role of examples at the end of the interview, Dr. Hickson said the following, which highlights the back and forth relationship that a mathematician may go through, even if he or she suspects that the conjecture may be true.

Dr. Hickson: You know the height a folly is to assume you know the answer when you don't. And so um, you might have a feeling or whatever that makes you spend more time one side of the problem. But, in my mind it's very important to not pretend you know something you don't. So you should always give some effort to both sides of the problem. And so I try to go back and forth. I'm like any other person though. I'm going to avoid work when I can. I'm going to pursue the easier option first… if I run out of gas then I'm going to reluctantly try the other one.

This back and forth dynamic provides some specific insight into the mathematicians' ways of viewing examples in the proving process in particular, but it is representative of a broader kind of metacognitive activity with which the mathematicians engaged. The mathematicians are able to recognize that a strategic choice of a counterexample or example could play a significant role in their understanding and proof of the conjecture. Other researchers (e.g., Savic, 2012) have indicated the importance of such metacognition among mathematicians. We suspect that this is a potentially important distinction between how mathematician mathematicians handle examples and how novice mathematics students handle them.

Conclusion and Implications

In this paper, we have highlighted particular ways in which work with examples manifest itself as mathematicians explore and prove conjectures. Our results point to the power of intentional example exploration in supporting one's understanding of conjectures and their proofs. Although the results presented here are based on a small set of interviews with mathematicians, the results are consistent both with the results from our large-scale survey of mathematicians' responses about their work with examples (Lockwood et al., 2012) and with aspects of existing literature (e.g., Harel & Sowder, 1998; Sandefur, et al., 2013; Weber, 2008). The interview data highlight the powerful role examples can play in exploring, understanding, and proving conjectures, as well as the ability to implement example-related activity in meaningful ways.

While we focus on mathematicians' work with examples and make no significant pedagogical claims, we cannot help but make some comments about how example-related activity we observed in experts might differ from students. The mathematicians' example-related activity stands in contrast to the role examples typically play in the work of mathematics students. The four themes listed represent aspects of mathematicians' example-related activity that could imply certain practices for students. For example, it might behoove students to explore the back and forth relationship between proving and disproving, and students might benefit from being exposed to various kind of examples in particular mathematical domains. And, too, pattern recognition or work with generic examples might be productive avenues that can lead students to successful proofs. Specifically, such activity might include having students answer reflective questions about their example choices as they explore a conjecture, or teaching them how a generic example differs from, but can lead to, a valid proof. A stronger understanding of the strategies mathematicians employ as they use examples to develop, explore, and prove conjectures may ultimately inform the design of instructional practices and curricula that effectively foster students' abilities to prove.

References

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

Antonini, S. (2006). Graduate students' processes in generating examples of mathematical objects. In J. Novotna, H. Moraova, M. Kratka, & N. Stehlikova (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 2, 57-64.

Ball, D., Hoyles, C., Jahnke, H., & Movshovitz-Hadar, N. (August, 2002). *The teaching of proof.* Paper presented at the International Congress of Mathematicians, Beijing, China.

Bills, L., & Watson, A. (2008). Editorial introduction. *Educational Studies in Mathematics*, *69*, 77-79.

Buchbinder O., & Zaslavsky O. (2009) A framework for understanding the status of examples in establishing the validity of mathematical statements. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education,* Vol. 2, 225–232.

Carlson, M & Bloom, I. (2005) The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58,* 45-75.

Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, DC: Council of Chief State School Officers.

Ellis, A. E., Lockwood, E., Williams, C. C. W., Dogan, M. F., & Knuth, E. (2012). Middle school students' example use in conjecture exploration and justification. In L.R. Van Zoest, J.J. Lo, & J.L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the Psychology of Mathematics Education* (Kalamazoo, MI).

Epstein, D., & Levy, S. (1995), Experimentation and proof in mathematics. *Notice of the AMS*, *42*(6), 670–674.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. *Proceedings of the 22nd Annual Meeting of the Psychology of Mathematics Education*, Vol. 2, 345-352.

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

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. *Issues in Mathematics Education*, *7*, 234-283.

Harel, G. (2002). 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* (pp. 185-212). Westport, CT: Ablex Publishing.

Harel, G. (2008). What is Mathematics? A Pedagogical Answer to a Philosophical Question. In R. B. Gold & R. Simons (Eds.), *Current Issues in the Philosophy of Mathematics From the Perspective of Mathematicians,* Mathematical American Association.

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

Iannone, P., Inglis, M., Mejia-Ramos, J. P., Simpson, A., & Weber, K. (2011). Does generating examples aid proof production? *Educational Studies in Mathematics, 77*, 1–14.

Kloosterman, P., & Lester, F. (2004). *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress*. Reston, VA: National Council of Teachers of Mathematics.

Knuth, E. (2002a). Proof as a tool for learning mathematics. *Mathematics Teacher*, *95*(7), 486–490.

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

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). New York, NY: Routledge.

Lockwood, E., Ellis, A.B., Dogan, M.F., Williams, C., & Knuth, E. (2012). A framework for mathematicians' example-related activity when exploring and proving mathematical conjectures. In L.R. Van Zoest, J.J. Lo, & J.L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 151 – 158). Kalamazoo, MI: Western Michigan University.

Lockwood, E. (2013). The strategy of solving smaller, similar problems in the context of combinatorial enumeration. Accepted with revision, *Research in Collegiate Mathematics Education*.

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

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics, 66*, 23-41. Doi: 10.1007/s10649-006-9057-x.

Pedemonte, B., & Buchbinder, O. (2012). Examining the role of examples in proving processes through a cognitive lens: The case of triangular numbers. *ZDM: The International Journal on Mathematics Education, 43*(2), 257-267. Doi: 10.1007/s11858-011-0311-z.

Polya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.
Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics, 21*, 589–598.
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.
Rittle-Johnson, B., Star, J.R., & Durkin, K. (2009). The importance of familiarity when comparing examples: Impact on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101*(4), 836-852.
Rowland, T. (2008) The purpose, design and use of examples in the teaching of elementary mathematics. *Educational Studies in Mathematics, 69*(2), 149-163.
Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics*. Doi: 10.1007/s10649-012-9459-x.
Savic, M. (2012). What do mathematicians do when they have a proving impasse? *In the Electronic Proceedings for the Fifteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education.* Portland, OR: Portland State University. February 23-25, 2012.
Schoenfeld, A. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education, 10*(3), 173-187
Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando, FL: Academic Press.
Simon, M. A. & Blume, G. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior, 15*, 3-31.
Sowder, L., & Harel, G. (1998). Types of students' justifications. *Mathematics Teacher, 91,* 670–675.
Stylianides, G. & Stylianides, J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*(3), 314-352.
Strauss, A. & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.
Tsamir, P., Tirosh, D., & Levenson, E. (2008). Intuitive non-examples: The case of triangles. *Educational Studies in Mathematics, 69*(2), 81-95.
Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples.* Mahwah: Erlbaum.
Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for research in Mathematics Education, 39*(4), 431-459.
Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12, 306-336.
Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to the Principles and Standards for School Mathematics* (pp. 227–236). Reston, VA: NCTM.
Zaslavsky, O., Nickerson, S. D., Stylianides, A. J., Kidron, I., & Winicki, G. (2012). The need for proof and proving: mathematical and pedagogical perspectives. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education: The 19th ICMI Study* (New ICMI Study Series, Vol. 15). New York: Springer.
Zazkis, R., & Leikin, R. (2008). Exemplifying definitions: A case of a square. *Educational Studies in Mathematics, 69*(2), 131-148.

STUDENTS' CONCEPTS IMAGES AND MEANINGS FOR AVERAGE AND AVERAGE RATE OF CHANGE

Eric Weber
Allison Dorko
Oregon State University

In this paper, we model students' concept images and meanings for average and for two- and three-dimensional average rate of change. We use these characterizations to describe how students use their meanings for average to interpret and reason about rate of change. We describe the importance of everyday meanings for average in students' conceptions of rate, and propose how instruction and activities might address this link. We conclude by discussing the significance of this work for mathematics education, and propose important directions for future research focus on students developing the meanings that instructors intend.

Keywords: rate of change, function, meaning, concept image, quantities

Background

The purpose of this paper is to characterize the role of students' meanings for average and the implications of the meanings for their understanding of average rate of change. To do so, we use students' concept images to model students' meanings for average, average rate of change, and instantaneous rate of change. Our focus on average rate of change takes into account students' meanings for both parts of the phrase: that is, their meanings for average and their meanings for rate of change. We argue that students construct understandings for average rate of change based on their meanings for average, and we discuss ways to use this tendency productively.

We think about rate of change as foundational to calculus because it allows a student to represent how fast a quantity changes with respect to one or more other quantities. We might imagine that a coherent understanding of rate of change relies on students imagining instantaneous rate of change as result of multiplicatively comparing the change of one quantity with respect to another over an infinitesimally small interval of change in the independent quantity. However, students do not always have the meanings we intend or think they have even though they use language similar to ours. Consider average as an example. The definitions of average and images associated with it come from everyday experience, language, statistics and calculus. These multiple meanings create what Barwell (2005) termed *lexical ambiguity*, which results when a common word is coopted by a technical domain. For instance, average can mean the arithmetic mean, the median, or the mode; it can also mean something 'typical' or 'normal.' Consider that the Japanese and Chinese words for average are translated as 'equal sharing', 'per-unit quantity', and 'smoothing out' (Cai, Lo, & Watanbe, 2002). The lexical ambiguity of average, and the variety of elements in concept images for it, may lead to situations in which students use meanings that might not be productive in calculus.

We hypothesized that students rely on images of average because researchers have found that students bring their everyday meanings for average to both mathematics and statistics (Kaplan, Fisher, & Rogness, 2009). These understandings of average range from average as a balance model (Hardiman, Well, & Pollatsek, 1984; Strauss & Bichler, 1988), to average as representative (Mokros & Russell, 1995), to average as typical (Kaplan et al., 2009). However, only one study (Cortina, Saldanha, & Thompson, 1999) has proposed a meaning for average (as a unit-rate) that seems productive for thinking about rate in calculus. No previous studies documented how students' interpretation of average affects student thinking about average rate and subsequently, rate of change "at a point". Given the prevalence with which students use their everyday meanings for average, it is likely that students interpret average rate of change in a way that leverages their everyday meaning for average. This suggests that many calculus students might rely on a meaning for average that does not allow them to think about average rate of change as a comparison of quantities varying, a meaning many instructors assume students possess. Exploring this avenue of research has potential to shape how we think about instruction about average in statistics and calculus, and more broadly, brings up the issue of attending to the meanings students have, in contrast to the meanings instructors assume exist for them.

Motivation for Study

Based on interviews with students in multivariable calculus, and our interpretations of the literature related to students' interpretation of mean and average, we hypothesized that students' understanding of average created incoherence for them as they learned about average rate of change in the calculus sequence. Specifically, we believed the meanings they constructed for average rate of change were based on their meanings (both everyday and mathematical) for average. We hypothesized that students relied on their meanings of the word average to interpret rate and derivative. We sought to characterize what students know about rate of change by focusing on their meaning for average and their subsequent ways of thinking about average and instantaneous rates of change.

In the following sections, we describe previous work regarding students' understanding of average, highlight the theoretical underpinnings of the study, identify how our assumptions about student thinking drove the study's design, illustrate our methodology and coding, and present a framework that characterizes students' concept images of average and average rate of change and the meaning we inferred from those images. We argue that specifically addressing the different uses of average in mathematics, statistics, and everyday language is crucial to students developing a coherent understanding of average and instantaneous rate of change in calculus. We close by suggesting a meaning for average that may help students develop the conceptions of rates of change that instructors intend.

Literature Review

Average Rate of Change

Students' thinking about rate of change as a measurement of how fast quantities are changing is foundational to calculus, yet many students possess difficulties reasoning about rate (Carlson, Larsen, & Jacobs, 2001; Rasmussen, 2000; Thompson & Silverman, 2008). Students' difficulties include problems interpreting the derivative on a graph (Asiala, Cottrill, Dubinsky, & Schwingendorf, 1997), and focusing on cosmetic features of a graph (Ellis, 2009). Thompson (1994) found that the difficulties students displayed in understanding the fundamental theorem arose from impoverished concepts of rate of change and incoherent images of functional covariation. Thompson (1994) described a coherent way of thinking about average rate of change of a quantity as, "if a quantity were to grow in measure at a constant rate of change with respect to a uniformly changing quantity, then we would end up with the same amount of change in the dependent quantity as actually occurred" (p. 239). We observed that this way of understanding was difficult for students to achieve.

The observation that Thompson's way of understanding rate of change is difficult for students to achieve is not novel. A number of researchers have proposed that difficulty might arise for students in understanding rate because of their inattention to quantities (Carlson, Smith, & Persson, 2003; Thompson & Thompson, 1994), from the students' association with rate as only a slope on a graph (Weber, 2012) or not understanding that a rate is a comparison of changes in quantities (Weber, 2012). However, no previous studies have suggested that students' understanding of average could affect their interpretation of average rate of change, and subsequently their way of thinking about derivative in both single and multivariable calculus. In the next section, we characterize the literature to suggest how the student's meaning for average could affect their understanding of rate.

Meanings for Average

Students' meanings for words used in technical domains are connected to past experiences with the word (Lemke, 1990). It is not surprising that students have various meanings for the word average, as the introduction and use of the word average varies across country, textbook, and classroom (Cai et al., 2002). Cai (2002) found that the meanings that textbooks promote for average are similar to "equal-sharing", and "per-unit quantity", but that these meanings are specific to textbooks in China and Japan. In contrast, United States textbooks tend to focus on the calculation to determine the average rather than students' interpretations of it (Cai, et al., 2002). The meanings Cai's study identified are consistent with the results from a number of researchers who have focused on characterizing students' understanding of average across mathematics and statistics (Kaplan et al., 2009).

Some early work about students' understanding of average identified students' "fair-share" or "balance" model of the mean (Hardiman et al., 1984; Pollatsek, Lima, & Well, 1981). By fair share, they meant a student thinking about the total magnitude being redistribtued across all elements such that each has an equal magnitude. Strauss (1988) described an example of the fair-share model using a balance beam in which values are placed an appropriate distance from the mean so the deviations from the mean are minimized. The mean as balance or fair-share model is consistent with the meaning of the word average from both Japanese and Chinese (Cai, 2002).

Mokros and Russell (1995) built on this work to propose that average is a tool for summarizing and describing a data set, and is thus context dependent (Mokros & Russell, 1995, p. 21). They interviewed 21 middle school students to characterize students' notions of average. They found that students' understandings of average include ideas like representative, reasonable or typical and that these understandings result from their everyday experiences. While they identified new categories for student thinking about average, we found that none of their categories were inconsistent with the fair-share model.

Consistent with the work in the 1980s and 1990s, Kaplan et al. (2009) found that students' meaning for *average* include ordinary, normal, typical, mediocre, not extraordinary, common, neither outstanding nor poor, standard, mean, median, in the middle, overall summary on something, general value that represents most of the data, overall outcome, mode, most common number, majority, a value we can use to compare one person's performance to the group, and the division of two statistics to get a whole answer. This group of studies suggests that both mathematics and statistics students have understandings of average that drew from colloquial language. It is also noteworthy that none of the studies identified an understanding of average as a unit rate, which is a foundational meaning for average in Chinese and Japanese textbooks.

Cortina et al. (1999) is the only study we know of that focused on students' understanding of average as a unit-rate that measures a group characteristic. These researchers suggested that characterizing the mean as typical of individual scores is vague because students may think the mean quantifies something about individual scores (Cortina et al., 1999, p. 1). They conducted teaching experiments with four middle school students as the students engaged in tasks intended to help them think about the mean multiplicatively. They found that their students' understanding of average was more than procedural and shared commonalities with Mokros and Russell's (1995) conception of "fair share". Furthermore, they determined even with the students' sophisticated understandings, it was difficult for the students to think about a mean as a measure of group performance. Cortina et al. (1999) concluded that it was key for students to 1) think of group performance as measurable attribute and 2) think about the mean as a measure of that attribute.

Together, these studies suggest that students' meanings for average are varied, reliant on context, focus on a characteristic of a group or set, and are often incoherent with each other. While none of the studies extend their students' understanding of average to thinking about average rate of change, they provide a basis for thinking about how a student might do so. For example, it is logical that a meaning for average as a unit rate would be productive for students thinking about average rate of change as a constant rate of change. It is also possible that thinking about average as a smoothing out or equal sharing could support a productive image of rate of change. However, it is also plausible that these ways of understanding average could result in misconceptions about rate, and thus lead to difficulty as students progress through calculus. This study contributed to filling this gap in research by focusing on characterizing the meanings students have.

Theoretical Framework

This section describes what we mean when we say we are interested in meaning, and how we studied meaning using students' concept images. Thompson (in press) traced the development of meaning as a construct throughout the 20th century to suggest that coherent meanings are at the heart of the mathematics that we want teachers to teach and what we want students to learn. Thompson built on Grice (1957) to argue that meanings reside in the minds of the person producing it and the person interpreting it (Thompson, in press, p. 4). As an extension of this assumption about where meanings reside, Thompson described how Dewey (1910) considered meaning and understanding as the product of thinking, and that coherence is an outcome of thinking (Thompson, in press, p. 5). This led Thompson to focus on what it means to understand. He relied on Piaget's notion of understanding as synonymous with assimilation to a scheme, where a scheme is an organization of operations and images. Thompson characterized Piaget's meaning for assimilation as similar to imbuing something with meaning, which goes beyond the standard description of assimilation as absorption of information. He concluded that constructing a meaning is similar to constructing an understanding, and that constructing a meaning occurs by repeatedly constructing understandings anew (Thompson, in press, p. 7). We drew from Thompson's characterization of meaning by focusing specifically on students' assimilations, in which they imbue meaning to something. We focus on the meaning they imbue to the word average and the phrase average rate of change.

We hypothesized that we could observe the products of students forming meaning by studying their images, definitions and representations for an idea. To do so, we relied on *concept images* (Vinner, 1983) as an orienting framework because we hypothesized that students' meanings for average were largely imagistic in nature or could be represented in that way. We drew on Vinner's (1983) definition of *concept image* as the set of properties associated with a concept together with the mental picture of the concept. However, we were not interested in using the characterization of a concept image as 'everything that comes to mind about a particular idea'. We were interested only in those elements of the concept image that helped us to infer the meanings that students had for average and average rate of change. Specifically, we wanted to understand the meaning that students imbued to these ideas by studying their concept images. We found Vinner's distinction between the concept image and the concept definition (the verbal definition typically used to introduce a concept) important because a concept requires both image and definition, "in thinking, almost always the concept image will be evoked" (Vinner, 1983, p. 293). That is, while engaging in mathematical thinking, students tend to use their mental pictures of a concept rather than a symbolic or verbal definition. However, they are only products of meaning, not meaning itself. Thus, to create models for students' meaning for average and average rate of change, we focused on their concept images.

Method

We hypothesized that students' understanding of average created incoherence for them as they learned about average rate of change in the calculus sequence, and they dealt with that confusion by relying on their everyday understanding of the word average to interpret rate and derivative. As a result, we sought to characterize what students know about average rate of change by focusing on their ways of thinking about average and their subsequent ways of thinking about average and instantaneous rate of change. We did so by creating tasks that would investigate the concept images students had for the word average and subsequent tasks that asked them to reflect on how their concept images and concept definitions affected their understanding of average rate of change.

Subjects and Setting

We interviewed sixteen multivariable calculus students from a pool of volunteers from four sections of multivariable calculus during the fall term at a large northwestern university. We chose this course because it was the students' first exposure to functions of more than one variable in mathematics. This allowed us to observe the students' initial fits and starts with systems with more than one quantity, and to adjust our subsequent questions to more clearly understand their thinking with particular regard to rate of change. Each student participated in a pre and post interviews occurred in the first and last two weeks of their course. The interview questions were designed to gain insight into students' understanding of function and rate of change. The pre-interview questions were open-ended and focused on single-variable functions and rates. The post-interview questions were also open-ended and consisted of questions about both single and multivariable rates of change, in addition to three tasks in which students were asked to compute the average of a given set of data.

Analytical Method

Data analysis was multi-phased. We used the pre-interviews to characterize students' understanding of function and rate of change. We identified common responses across interviews using grounded theory (Corbin & Strauss, 2008) and characterized students' concept images for average, average rate of change, and instantaneous rate of change. Our analyses from the pre-interviews suggested that some students relied on definitions of average in their representations of average rate of change, and that those definitions were prevalent in both two and three dimensions. We designed the post-interviews to gain insight into students' meanings for average,

average rate of change, and their thoughts about how those uses of average were related. We constructed a set of concept images for both average and average rate of change using open and axial coding (Figures 1 and 2) to describe a model for the meanings we believed students had.

Results

Concept Images

Figures 1 and 2 represent the outcome of open and axial coding of students' concept images for average and average rate of change. Each individual word in the category column resulted from the open coding of students responses during the interviews. During process of axial coding, we determined that a number of categories were representative of a similar concept image, and we grouped them together to create five categories for *average* and four categories for *average rate of change*. We created specific criteria for an instance of an interview to be coded as a particular category, and used examples from the data to demonstrate what we meant by particular criteria for other coders (inter-rater reliability K> 0.70). These categories provided the basis for describing students' meanings for average and average rate of change, often involving multiple categories for an individual student.

Category	**Criteria**
Normal, typical, mediocre, common	Student uses the word 'normal,' 'mediocre,' 'typical,' or 'common' to describe 'average.'
Mean	Student uses the word 'mean' to describe 'average' or as a synonym for 'average'
Median, middle, center, balance point	Student uses the words 'median,' 'middle,' 'center,' or 'balance point' or talks about the average as being the middle or center of the data
Overall summary, representative value, value used to compare, estimate, expected value	Students talk about the average as a number that presents an overall summary of the data; a number that is representative of all the data; the average as an estimate/approximation or expected value for a new data point; or talk about using the average to compare data
Mode, most common number	Student uses the word 'mode' as synonymous for 'average' or talks about average as the most common number
Smoothed-out value	Student talks about the average as a smoothing out of all the values or redistributing the data equally across *n* pieces
Unit rate	Student discusses average the constant rate at which a quantity would need to change to produce the same overall change that was originally observed

Figure 1: Students' concept images for average

Category	**Criteria**
Arithmetic mean of slopes	Student talks about summing slopes and dividing by number of slopes summed whether it is a finite or infinite number of slopes is irrelevant to the student.
Expected or most common slope	Student uses the word ‘most common’, ‘expected’, ‘typical’ to describe average rate of change. The student expects the average rate of change provides information about ‘all’ of the slopes.
Representative slope or rate of change	Student describes the average rate of change as representative of all the rates or slopes. May describe providing information about each of them.
Constant rate of change	Student describes average rate of change as the constant rate of change required to produce the same change in the function over the original interval of input.
Smoothing out of all the slopes	Student describes the average rate of change as the all of the slopes smoothed out. Student describes decreasing the ‘choppiness’ of the slopes.

Figure 2: Students’ concept images for average rate of change

Example of Coding for Concept Images

In this section, we present excerpts to illustrate the coding process. The following results are representative of our findings for the sixteen students, which we expand on further in subsequent results. We have included students’ responses to highlight major categories of the concept images we identified. We share our coding and inferences by focusing on excerpts from student responses to the following three tasks:

(1) The data given below represent the masses of six fishing lures. What would the average mass of the lures mean?

(2) Suppose we define a function *f*, so that $f(x)=e^{-\cos(2^x)}$. Discuss the *process* you would use to determine the average rate of change of the function with respect to *x* over the interval [2.0, 2.2].

(3) Suppose we define a function f, so that $f(x,y)=e^{-\cos(xy)}$. Discuss the *process* you would use to determine the average rate of change of the function. What information do you need to know to complete this process?

Figure 3: Selected Interview Tasks

These tasks are representative of the semi-structured tasks that we presented. We asked students this sequence of questions in order to have them describe average and average rate of change within five interview minutes. We then asked them to reflect on the similarities and differences in the use of average in each question. Our analysis of students’ responses allowed us to construct descriptions of their concept images and to make inferences about their meanings. Below, we present three students' responses to the previous tasks to demonstrate our coding of students' discussions.

Student 1

Brian: *[Response to 1] I see the average as kind of like adding everything up into a big ball, and then smoothing it out into equivalent pieces.*

[Response to 2] I see the average rate of change like a constant rate of change. Like, how fast the function would need to change to produce the same change in y over the same change in x, but at a constant rate. You take the change in y over the change in x, that kind of smooths it out to determine it for you.

[Response to 3] Now, well, this is harder but I still know I am finding a constant rate of change. However, to pick a constant rate, you have to specify a direction in space, or there would be infinite average rates of change. So, you still have a change in the function on top, but divided by a change one variable or the other. It tells you a constant rate of change.

We characterized Brian's concept image for average as a "smoothing out" of all the values in the data set or the function. A number of students' concept images included this idea, in particular as a response to the average mass of a lure. These students discussed the average mass as what one would find it one melted all the lures into one piece of metal and divided it into six equal parts. Brian used "smoothing out" to discuss the average rate of change as the constant rate at which the function would need to change to produce the same change in y over the same change in *x*. His idea of smoothing out for a function seemed to be the function "smoothed" into a secant line. Brian thus used his 'smoothing out' idea productively because he could connect it to variation in quantities and, we inferred, a secant line. It is telling that he did not use the phrase 'smoothing out' in talking about rates of change in three dimensions; rather, he discussed the constant rate of change in a particular direction. Thus, while his verbal description of average and average rate of change seemed to be focus on smoothing out, he still retained a meaning for rate as a comparison of changes in quantities.

Student 2

Jordan: *[Response to 1] Well, I sum the masses, then divide by how many there are, which tells me what their mass was mostly, or typically.*

[Response to 2] Well, I am finding the slope between two points here Right, so I find the change in y over the change in x. That just tells me a typical slope.

[Response to 3] Again, I probably am finding a slope, an average slope, so I need a change in something over a change in something else. Probably a combination of z, x and y? Again, it would just tell me a typical slope.

We characterized Jordan's concept image as of a "typical" value, an idea that prevented her from thinking about average rate of change as the constant rate of change needed to have the same change in y over change in *x*. In other words, 'typical' prevented her from thinking about rate of change *as a rate* as instructors might expect. However, Jordan appeared think about quantities changing, as evidenced by her comments about a change in *y* over change in *x*, but her meaning for comparison of these changes did not extend beyond finding a change in a quantity to divide by a change in another quantity. We believed this for two reasons. First, it seemed that she was reciting a definition of slope that had meaning in terms of computation – "change in y over change in x" seemed to be a command to calculate that would tell her about a typical change. Second, when she tried to extend finding the changes in y over the change in x to three space, she concluded the denominator would be "some combination of x, y, and z."

We constructed concept images and coded them in similar ways for each of the sixteen interview participants. After we coded for concept images, we noted that many students appeared to hold multiple concept images of average in mind, and applied those understandings depending on the context of a particular task. This suggested to us that students' meanings were flexible and resulted in the student imbuing the phrase "average" with a particular meaning depending on the elements of the task. In the next section, we provide a broader overview of our results by describing how we used students' concept images to make inferences about the meanings they possessed.

Representative Excerpts

Mathematically, average may refer to mean, median, or mode and colloquially it may refer to a value that is typical, common, representative, or expected (among others). It was not surprising that students' concept images reflected these multiple meanings. We gained insight into students' concept images by introducing tasks that pushed them to describe things with which they were unfamiliar. For example, we asked them to describe the meaning of average void of a particular context. Many students expressed difficulty with explicitly defining average ("everyone knows what average means"), and we think that forcing them to search for words with which to define it made for particularly clear snapshots of their concept images.

We noticed in our analyses that while many students held multiple understandings of average in mind simultaneously and without incoherence, the meanings they appeared to possess for average were robust across tasks. By robust, we mean that they found creative ways to apply their understanding of average to interpret both average and instantaneous rate of change in two and three dimensions. By extending this meaning for average to multiple situations and tasks, we believed that the students were seeking coherence in their understanding. They appeared to think that if average meant the same thing in each context, then their understanding of average was accurate. Thus, their desire for coherence resulted in them attempting to apply similar meanings for average across problem situations. In the following table we present an overview of student responses that demonstrate the coherence they created for thinking about average and average rate of change.

Key Issues

The selected excerpts suggested to us that students' meanings for average centered on a measure of central tendency, typicality, or representativeness. The data also supports that students think about average rate of change as a measure of central tendency of the rates or slopes. More broadly, they think about average and average rate of change as the measure of a characteristic of a group that also provides information about the individual observations in that group. While we believe that these meanings for average can be leveraged in the context of instruction, our interpretations of the data suggest that the students' meanings allow them to believe their understanding is coherent, even when an expert may seem these as inaccurate.

STUDENTS' KNOWLEDGE RESOURCES ABOUT THE TEMPORAL ORDER OF DELTA AND EPSILON

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

The formal definition of a limit, or the epsilon delta definition is a critical topic in calculus for mathematics majors' development and the first chance for students to engage with formal mathematics. Research has documented that the formal definition is a roadblock for most students but has de-emphasized the productive role of their prior knowledge and sense making processes. This study investigates the range of knowledge resources included in calculus students' prior knowledge about the relationship between δ and ε within the definition. diSessa's Knowledge in Pieces provides a framework to explore in detail the structure of students' prior knowledge and their role in learning the topic.

Key words: limit, formal definition, students' prior knowledge, fine-grained analysis

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

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

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

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

Although studies have sufficiently documented that the formal definition is a roadblock for most students, little is known about how students actually attempt to make sense of the topic, or about the details of their difficulties. Most studies have not prioritized students' sense making

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

processes and the productive role of their prior knowledge (Davis & Vinner, 1986; Przenioslo, 2004; Williams, 2001). This may explain why they reported minimal success with their instructional approaches (Davis & Vinner, 1986; Tall & Vinner, 1981). Thus, understanding the difficulty in the teaching and learning of the formal definition warrants a closer look – with a focus on student cognition and with attention to students' prior knowledge. It also calls for a theoretical and analytical framework that focuses on understanding the nature and role of students' intuitive knowledge in the process of learning.

A small subset of the studies have begun exploring more specifically student understanding of the formal definition (Boester, 2008; Knapp and Oehrtman, 2005; Roh, 2009; Swinyard, 2011). They suggest that students' understanding of a crucial relationship between two quantities, ε and δ within the formal definition warrants further investigation. Davis and Vinner (1986) call it the *temporal order* between ε and δ, that is the sequential ordering of ε and δ within the formal definition where ε comes first, then δ [2] (p. 295). They found that students often neglect its important role. Swinyard (2011) found that the relationship between the two quantities is one of the most challenging aspects of the formal definition for students. Knapp and Oehrtman (2005) and Roh (2009) document this difficulty for advanced calculus students. This difficulty is also prevalent among the majority of calculus students who struggled with the formal definition in Boester (2008).

One hypothesis in the literature for the difficulty is that to understand the temporal order of ε and δ requires the act of "reversing the function process" (Oehrtman, Carlson & Thompson, 2008). Swinyard (2011) calls it the "y-first" conception because ε, the error bound for the output y comes first. This act contrasts with the usual functional relationship that most students are familiar with, that is input, x first, then output, y. While Swinyard (2011) shows that such step is crucial in understanding the definition, *how* students arrive at that understanding or begin to reverse the function process remains an empirical question.

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

1. What claims do students make about the temporal order of ε and δ?
2. If students in fact struggle with the temporal order, what is the nature of their difficulty?

To explore the nature of the difficulty this study will explore the range of *knowledge resources* students use to make sense of the temporal order. *Knowledge resources* are defined as relevant prior knowledge that might be used to reason and justify a claim. In this study we focus on *observable knowledge resources*, that is knowledge resources that can be inferred through students' assertions, gestures and artifacts. For brevity we will use knowledge resources for the remainder of the document when we talk about the observable knowledge resources.

Theoretical Framework

[2] Davis and Vinner (1986) used the phrase *temporal order* to describe the sequential relationship between ε and N in the formal definition of a limit of a sequence. Knapp and Oerhtman (2005) have used the same phrase to describe the relationship between ε and δ as they follow the same sequential ordering.

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

Methods

The data for this study comes from two rounds of pilot study for a larger study investigating the role of prior knowledge in student understanding of the temporal order. The main source of data is individual student interviews with seven calculus students. The goal of providing a detailed account of knowledge structures and learning processes suggests the use of a relatively small number of research subjects, given the depth and detail of analysis that will be offered. Each of these students has received some form of instruction on the formal definition, so we anticipate some knowledge about the definition to be a part of their prior knowledge.

The interview protocol consisted of a task portion to re-familiarize the student with limit and then an opportunity for the student to explain and define limit in their own words. Next, both general and specific questions were asked about the parts of the formal definition. Afterwards, students were asked specific questions about the temporal order of ε and δ. To explore the stability of students' knowledge across different contexts, we asked students about the temporal order of the two variables in four different contexts: dependence, control, time, and ordering of the four variables in the definition (for the questions see Table 1 under results). In those contexts, the relationship between δ and ε could be described as: δ depending on ε, one is trying to control x using δ, ε comes first, and an ordering of the four variables in which ε comes first before δ. There are multiple ways to order the four variables within the definition. One might argue that ε comes first and δ is determined from ε, then one uses the x values that are within δ of a to see if they yield function values that are within ε of L. So one ordering could be ε, δ, x and $f(x)$. Each individual interview lasted about 2 to 3 hours. These interviews were videotaped following recommendations in Derry et al. (2010).

Analysis

The first part of the analysis places students in categories based on their claim about the temporal order of ε and δ. There will be three categories: *epsilon first*, *delta first*, and *no order*. For a student to be classified into the category *epsilon first*, s/he would respond in the following way to the four questions. S/he would say that: 1) δ depends on ε; 2) one is trying to control x using δ, based on a given ε; 3) ε comes first and then δ; 4) the four variables are ordered in such a way where ε comes first then δ. For a student to be classified into the category *delta first*, s/he would respond in the following way to the five questions. S/he would say that: 1) ε depends on δ; 2) one is trying to control $f(x)$ using ε, based on δ; 3) δ comes first and then ε; 4) the four variables are ordered in such a way where δ comes first then ε. For a student to be classified as *no order*, there needs to be variance in responses across the different questions. In this study, we found few inconsistencies between the four different ways of asking the question with these seven students but we expect more variance with a larger population of students.

The second part of the analysis explores the range and nature of knowledge resources. As stated earlier, we operationalize *knowledge resources* as relevant prior knowledge that might be used to reason and justify a claim that can be inferred through students' assertions, gestures and artifacts. We identify knowledge based on what students say in the moment using the Knowledge Analysis (KA), which is a methodology consistent with KiP (Campbell, 2011; diSessa, 1993; Parnafes and diSessa, in press; Schoenfeld, Smith & Arcavi, 1993). KA is a combination of top-down and bottom-up qualitative analysis. This means that emergent knowledge structures must be held accountable to both empirical evidence and existing theories and literature. The analysis focuses on discussions around the four questions about the temporal order of δ. To hold our conclusions accountable to the data, reasonable interpretations for the statement will be considered and be put through the process of *competitive argumentation* (Schoenfeld et al., 1993) using other parts of the transcripts. With each of the observable knowledge resources, particular care will be given to investigate their origin and when it originally came up. Until there is consistent evidence of stance taken by a student, it would be impossible to make claims about the stability or how committed the student might be to the specific claim they made.

Results

Relationship Between the ε and δ

Five of the seven students we interviewed concluded that δ came first, 2 students concluded that ε came first and no student fell into the no order category. These findings are consistent with the literature's claim that students struggle with the temporal order within the formal definition. The table below shows the claims students made about the temporal order between ε and δ across the different contexts. We determine the student's final categorization by what the student said last about the relationship between ε and δ.

Table 1
Students Responses Across the Different Contexts of the Temporal Order

Student Initial	Question 15: Which depends on which?	Question 17: Which variable are you trying to control?	Question 18: Which one comes first?	Question 19: Order of [the four] variables.	Final categorization
DC	δ depends on ε	[Skipped]	N/A	N/A	Epsilon first

DL	ε depends on δ	δ you can control, ε you're trying to control.	N/A	N/A	Delta first
JJ	ε depends on δ	Control δ and [trying to] control ε.	N/A	N/A	Delta first
AD	δ depends on ε	[Skipped]	ε is first, you break down the epsilon to find delta.	Student decided not to try.	Epsilon first
DR	ε depends on δ	Trying to control δ so that you can get a smaller ε.	You get delta first then you get ε as your result.	$x, f(x), a, L, \delta, \varepsilon$	Delta first
SR	ε depends on δ	Trying to control x and $f(x)$, but not sure.	Calculate delta first, then used to calculate ε.	$x,\ \delta, f(x),\ \varepsilon$	Delta first
OB	ε depends on δ	Try to control x and δ to find ε.	Delta comes first, ε second.	$\delta, x, a, f(x),\ \varepsilon$	Delta first

The five students who ultimately concluded that δ was first were consistent in their inference of the relationship between the two variables in the other questions. All five concluded the temporal order by claiming that ε depended on δ. Four of the five (DL, JJ, DR and OB) believed that one could control δ in order to control ε. Three of those five students (DR, SR and OB) were part of the second round of pilot and their responses to the other order question (question 19) were also consistent. While one (SR) was unsure whether one would calculate ε as a result, the ordering is consistent in that δ comes first and get ε second. SR was also unsure of which variables they were trying to control. It's worth noting that even the two students who ended up with the category *epsilon first* initially thought that *delta first* made more sense to them. We will see some of their explanations in the next section. Furthermore, due to time constraints and different versions of the protocol as it developed, some questions were skipped.

Range and Diversity of Knowledge Resources

The following diagram provides an overview of the knowledge resources that emerged from the data. We inferred them from the students' explanations about the temporal order across the different contexts and organized them into two categories: those that pertain to the formal definition and those that did not. Each section below will provide an illustration of some of the common knowledge resources across different students. Most of the resources will support the claim that ε comes first; though some will support the normative claim that δ comes first. We start with the most common knowledge resource, then we continue in order from left to right on the diagram.

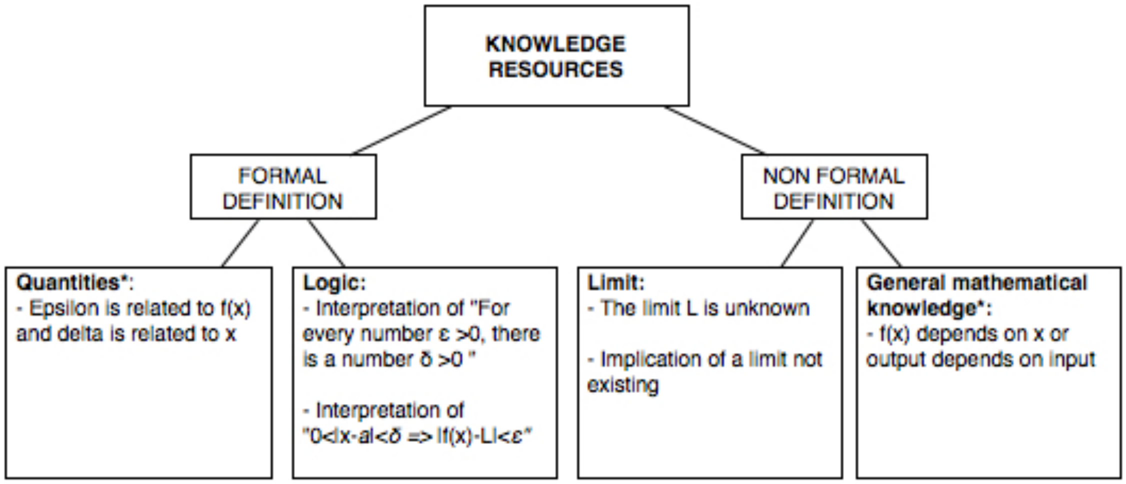


Figure 1. Six documented knowledge resources from the seven students. Star indicates knowledge resources that were often cued together.

The "$f(x)$ depends on x" and "δ is related to x and ε is related to $f(x)$" Knowledge Resources

One very common knowledge resource that emerged in the pilot study was the *output $f(x)$ was dependent on the input x.* In fact, all five students who ended up in the *delta first* category used this particular resource. The two students, who concluded that ε was first, also used this resource at some point in their explanation.

Students often associated this knowledge resource with another knowledge resource that ε was a quantity related to $f(x)$ and δ was a quantity related to x. With this combination of knowledge resources, students often concluded that $f(x)$ depended on x meant that ε must also depend on δ, and thus δ was first. In our data, we found two ways that students used this knowledge. One of the ways was students argued that δ was associated with x and ε was associated with $f(x)$ so since $f(x)$ depended on x, then ε depended on δ. The second way was that students said that δ was either influencing the input or *was* the input and ε dealt with the output or *was* the output. Furthermore, since output depended on input then ε depended on δ. For example, DC illustrates the first way below.

> Um [inaudible] well given that the, um, delta does generally or does seem to refer to the x value or the range of x values, the domain of x values that you want to be paying attention to, generally I think of functions, um, since a function is a relationship between dependent and independent variables, I tend to think of x as being you know as they are the, uh, independent variables. And so the y as being the ones that are altered by the x. So that's how you plug in numbers for functions, that's how you utilize functions in most cases. So it makes more sense to me to think that as epsilon being dependent on delta, where I'm assuming that delta is referring to x and epsilon is referring to y values" (turns 137-145).

DC reasoned that δ referred to a range of x-values and thus ε referred to a range of y-values, and since $f(x)$ or y depended on x, then it 'makes more sense to him' that ε depended on δ. In this case we would argue that DC used the following knowledge resources to conclude that δ was first: 1) the dependence between x and $f(x)$; 2) δ refers to x values; and 3) ε refers to y values. Observe the similarity between what DC said with what DR said in her interview: "Um, see cus I was looking at it like the x or the $f(x)$ or the yeah, the $f(x)$ depends on the x and that's how I was like saying that epsilon depends on delta because epsilon like is related to the $f(x)$ or whatever" (turn 578). DR also relied on the dependence between x and $f(x)$, and she, too saw ε as "related to the $f(x)$." Note that DC was one of the students who ended up in the *epsilon first* category.

DL and JJ also reasoned about the formal definition using the knowledge resource of x was associated with δ and $f(x)$ was associated with ε. DL concluded that ε depended on δ because "[epsilon] is the $f(x)$ and $f(x)$ depends on x that epsilon would depend on delta" (turn 170). From his response it seems that he also had equated ε with $f(x)$. For JJ this way of thinking was revealed when he discussed question 17, which of δ and ε he had control over and which he was trying to control. He argued that he had control over δ to control ε because, "Cus of x and y. I equated delta with being x and epsilon y" (turn 260). In sum, four students (DL, JJ, DC, DR), either understood ε and δ to be y and x values respectively or at least associated with them. They then used the dependence between x and y (or $f(x)$) to conclude the dependence between ε and δ.

On the other hand, AD did not explicitly relate δ to x and ε to $f(x)$, but he still used the input/output relationship between the function and its domain to make conclusions about the dependence between ε and δ. He understood δ to be a bound that restricted the closeness between x and a. However, for AD the process started with δ restricting the input and seeing what happened with the output. Below is his response for why ε depended on δ.

310 AD … [I]t's because delta is um, you're saying delta must be greater than the input minu-subtracted by what you're centered around.

311 AA Uh-hm.

312 AD So you're saying that delta must be, the interval around a number *a* must be less than delta, so you're saying, um the input cannot get outside of this region, it cannot be getting, [inaudible] our region, this interval it, it cannot get exceedingly big.

313 AA Uh-hm.

314 AD And then for epsilon you're evaluating x around *a,*

315 AA Uh-hm.

316 AD and then you're subtracting 1. When you plug in x for, when you plug in *a* for x, so what winds up happening is you're seeing how big the difference [of the function value] is between a number near *a* and then *a* itself.

317 AA Uh-hm. So how does that say that epsilon depends on delta?

318 AD It's, uh, because your input, your delta is influencing your input and then epsilon must be greater than your input minus your input of *a,*

319 AA ok, your output [correcting]

320 AD your output [in agreement].

321 AA And since, so since output depends on input..

322 AD Yes.

323 AA epsilon depends on delta…

324 AD Yes.

Here, unlike the four other students, AD did not specifically refer to the dependence between x and $f(x)$, instead he was referring to the relationship between input and output. While the wording was different, we took these two to represent the same knowledge resource. Moreover, AD might have had a slightly different conception of δ and ε. AD stated that δ was not just a quantity related to *x,* but more specifically, it was "influencing your input" (318). There was a sense of δ constraining the input that the other two students' interpretation of δ did not have. Furthermore, ε was also stated as something more than just dealing with *y* values or $f(x)$. AD stated that ε must be greater than your [output] minus your [output] at *a*. So we see here while AD might have interpreted ε and δ differently than DC, JJ, DL and DR, AD still cued the same knowledge resource, "the dependence between *x* and $f(x)$."

Interpretation of "For every number ε >0, there is a number δ >0" as Knowledge Resource

Another knowledge resource that was used by four of the seven students was interpretation of part of the statement of the formal definition "for every number $\varepsilon > 0$, there is a number $\delta > 0$." For two of the students, this statement was used to justify that δ comes first. We see this in JJ's interview when he says "If for every number epsilon, there is a number delta sounds like epsilon is based on, if there is a number delta or not" (turn 232). In this excerpt, JJ reasons that ε is based on the existence of δ; in other words, δ comes first. Similarly, DL interprets the statement to mean that δ had to exist in order for an ε greater than zero to exist when he says, "It seems that delta is supposed to affect ε, for me. Um, but on here [the statement of the definition] it seems that, that epsilon is affecting the delta, the way that I'm reasoning it, is that, since it exists, that epsilon is greater than zero, there must have existed some delta to have made this epsilon greater than zero" (turn 59). Notice that DL recognized that the statement of the formal definition suggested that δ depended on ε, or ε comes first, but his reasoning used this part of the statement and he concluded the opposite relationship.

In contrast, a third student, DC, uses his interpretation of the statement to support the claim that ε comes first. He said, "Actually the way I should be saying that, it's actually said in the form of this text. That for every epsilon there is a delta. Not for every delta there is an epsilon such that these conditions are met. So delta is dependent upon epsilon. It seems" (turn 123). He later confirmed that opinion and said, "Because as the, um, definition says, uh it's for every number epsilon, there is a number delta. So your delta is dependent on epsilon." (turn 196).

Interpretation of the Statement "If $0<|x-a|<\delta$ *then* $|f(x)-L|<\varepsilon$" *as Knowledge Resource*

Three students (JJ, SR and AD) used this part of the statement to conclude that δ comes first in the temporal order. To JJ this part of the formal definition said that a statement about δ implies a statement about ε, so ε depended on δ. He said, "Because if, if this [*points at* $0|<x-a|<\delta$] is true then that's [$|f(x)$-L$|<\varepsilon$] true. Um, and this has to do with delta and this has to do epsilon" (turn 218).

SR noted that this statement indicated that with the formal definition one would start with the *x* value. SR argued that the *x* had to be within a certain δ before the ε can be satisfied, and she used that as a part of her reason for why ε depended on δ. SR argued,

> My reasoning is that um, my reasoning is that you have to first /…/ look at /…/ the *x* value that you're approaching and it has to satisfy these conditions and be within a certain delta for there to even be /…/ a /…/ limit, /…/ like your whole deal with the epsilon, it [delta] has to be satisfied first, so that's why I feel like it depends on delta (turn 319).

SR later said that she got this idea from the if-then part of the formal definition and confirms that, to her, the if-then statement is telling her that the δ is "there first" (turn 327).

The "limit L is unknown" Knowledge Resource

One student, OB used their knowledge about limit more generally to reject that δ depended on ε. Specifically, he believed that in the context of a formal definition one still would not know what the limit was. That is one would try to find the limit and would do so by evaluating the function for values of x near a. That is, one would start with values of x near a then find what happens to the value of $f(x)$. The implication was that one would start with δ then compute ε.

OB argued that one would start with δ. Below he explained why one would not be starting with ε instead. He argued,

> Let's just say the contrary. If we had chosen epsilon to be the first number. Then it would mean that we would, we would have to know L. So it's pretty much just going around, if we know L then what's the purpose of trying to find an epsilon? Whereas if we first choose delta, we already know x and a. Uh and then we would find the epsilon so delta comes first (turns 144-152).

OB was convinced that within the context of the formal definition one would not know what the limit was. To him this means that one would find ε in order to figure out what the limit is. In this case we see two students who believed that the limit was still unknown in the context of the formal definition, much like calculating limits usually, and thus ε was used to get close to L, so δ comes first.

Implication of the Limit Not Existing as a Knowledge Resource

Another knowledge resource that students used to reason about the temporal order between δ and ε is what would happen when the limit did not exist. Two students, DC and AD used the context of a limit not existing to infer the dependence between ε and δ. DC came up with a problem context where the limit did not exist where there was a vertical asymptote at a and the left hand limit does not equal the right hand limit.

> [I]f you have a divergent asymptote then that means your epsilon will be infinitely large. So you can't really. It seems that you can't make any statement at all about delta in that situation… if you can't, um, like really confine that [ε] then you can't have any kind of meaningful knowledge about delta, it seems to me. Because as the, um, definition says, uh it's for every number epsilon, there is a number delta. So your delta is dependent on epsilon, and if you can't um like really confine that [ε] then you can't have any kind of meaningful knowledge about delta, it seems to me (turn 196-198).

DC argued that in this problem context ε would be infinitely large which would deem delta useless in confining x around a to confine $f(x)$ around L. As such depending on what ε is, δ might be meaningless, hence δ depends on ε, which is what DC ultimately concluded.

AD referred to the same problem context but only went as far as saying that δ and ε depended on each other, albeit being a weak relationship. He argued that in the case when the limit did not exist, then the δ being small would not necessarily mean that ε would also be small. He argued,

> So, um, they- they depend on each other a little but not like completely. It's, a weak I'd say it's more of a weak, um, connection because, uh if the limit exists then there's gonna be some sort of you know, if, as this [*possibly points* at $0<|x-a|< \delta$] gets you know smaller this [$|f(x)-L|< \varepsilon$] is getting, the difference is gonna get smaller. But if the limit doesn't exist like it did in 4 [problem 4 with diverging limits], where it, it, you know, it doesn't exist then as this [*points* at $0<|x-a|< \delta$] gets smaller this [$|f(x)-L|< \varepsilon$] isn't gonna change, this isn't gonna help [inaudible]. So it's gonna, still not exist even though the function is getting close [*gestures his two hand coming together*] to that point (turns 292-296).

AD seemed to be arguing that when the limit existed then as the δ expression ($0<|x-a|< \delta$) got smaller then the ε expression ($|f(x)-L|< \varepsilon$) would get smaller as well but sometimes, as in the case when the limit does not exist, the δ getting smaller would not affect the ε. But instead of concluding that then δ would be dependent on ε like DC, AD concluded that in that case there would be a 'weak connection' between δ and ε.

Summary

We have shown the diversity and range of knowledge resources students used to infer the temporal order of δ and ε. We started with the most common resource of *"$f(x)$ depends on x"* and *"δ is related to x and ε is related to $f(x)$."* In fact, all students who concluded that δ was first used this resource. We also found particular interpretations of the two parts of the statement of the definition, and also "the limit L is unknown" and the implications of the limit not existing. We grouped these resources into two categories, those related to the formal definition and those who did not. We also saw how some of these knowledge resources were cued together. For example, students' understanding of the quantities ε and δ were cued along with the dependence relationship between x and $f(x)$ in inferring the relationship between δ and ε, like with DC and DR.

Discussion and Implications

The result of the first part of our analysis confirms the literature's claim that the temporal order of ε and δ is difficult for students to understand. The majority of the students in our study concluded that δ came first, and even those who finally concluded that ε came first at some point during the interview stated that δ being first made more sense to them. The result from the second part of our analysis began to uncover the complexity of the issue. The second part of our analysis explores the diversity and range of knowledge resources students use to make claims about the temporal order. We showed that students arrived at their conclusion using a variety of resources, from the most common combination of resource of "$f(x)$ depends on x" and "δ is related to x and ε is related to $f(x)$," to the less common "the limit is unknown" with the one student OB. We also found some knowledge resources that assisted students to conclude that ε came first. "Implications of the limit not existing" and the normative interpretation of the statement "for every number $\varepsilon > 0$, there is a number $\delta > 0$" were both knowledge resources that assisted AD and DC to conclude that ε came first.

In some ways, the fact that "$f(x)$ depends on x" is the most common knowledge resource that conflicted with the temporal order of δ and ε is not surprising. Students are used to seeing functions as going from x to $f(x)$, so $f(x)$ depending on x, and the temporal order just follows the “opposite” direction, much like what Oehrtman et al. (2008) suggested with the reversal of function process. But our result elaborates on that conflict. This particular resource on its own is not necessarily problematic, but when it is cued with a particular understanding of δ and ε, that is “δ is related to x” and “ε is related to $f(x)$,” it led many students to conclude that ε depended on δ. The natural question is why did so many students have this “gloss” of ε and δ? What we found in our data was that most of them relied on the syntax of the definition to make meaning of the two variables. Specifically, many inferred from the statement “if $0<|x-a|<\delta$ then $|f(x)-L|<\varepsilon$” that “δ is related to x” and “ε is related to $f(x)$,” often ignoring the “less than” symbol. This tendency relates to our next point.

Notice that the resources are all mathematical inferences and mathematical interpretations, and there were no intuitive ideas. For example there were no ideas about approximation or error bounding like ones used in Oehrtman (2008). There are two possible explanations for this. First, it might be that this shows the lack of access into the formal definition for students. That is, the statement of the definition as presented in most textbooks provides little space for students to capitalize on their intuitive knowledge. As a result, many of them only had the syntax (the symbols and the logical structure) of the formal definition to make sense of it. And as we saw in our study, students came up with their best interpretation and often it led them to the wrong conclusion.

Second, the prevalence of mathematical inferences and interpretations might be a limitation to our current analysis. That is, we have not gone small enough in grain size of knowledge or specific enough to uncover students’ intuitive knowledge. For example, it might be that the “gloss” for δ and ε is indicative of a typical pattern in explaining. That is, while students might have a more particular view of δ and ε, in explaining the temporal order they might feel that a “gloss” is sufficient, especially when the gloss would be consistent with their understanding of how functions usually works. We have evidence that the student AD might be doing just this in developing his claim. This is a hypothesis that we are currently exploring in the larger study about students’ intuitive knowledge about the formal definition.

Our results show the importance of understanding students’ prior knowledge in understanding the learning of the formal definition. We have discussed the implication of the most common knowledge resource, but from our analysis we also found other knowledge resources that supported the claim that ε depended on δ, or δ first. It is important for us to reiterate that these knowledge resources are not misconceptions. For example, no one will say that the dependence of $f(x)$ on x is a misconception. For us, the importance lies in understanding how these resources are cued, how they might interact with others to lead to whichever conclusion about the temporal order.

The question for practice is then how can we begin to address the issue knowing the diversity and range of the knowledge resources students use to reason about the temporal order? Following the recommendations in Smith et al. (1993), we cannot expect to *replace* these seemingly unproductive knowledge resources. Instead, in instruction we need to figure out different ways for students to reorganize these resources, and to help them understand the appropriate contexts to use them. For example, in the larger study, we are attempting to take advantage of students’ intuitive knowledge about idea of *quality control*, to not only assist them in understanding the formal definition, but also to help them distinguish the difference between error and error bound, that is the difference between $|f(x)-L|$ and ε and in turn also help them distinguish the dependence between the errors and the error bounds. We have some evidence for the productivity of this approach.

In conclusion, while we have yet to see other patterns that will be revealed from interviewing more students, our findings so far suggest that students reason about the formal definition in particular ways. There is a range of knowledge resources students use to do so, and we expect to find more as we interview more students. It remains to be seen if we will find more mathematical interpretations or uncover more intuitive knowledge that students use in reasoning about the temporal order. Either way, we believe that a better understanding of the nature of these resources can facilitate the design of instruction that can help students bridge and reorganize these resources for a better understanding of the formal definition.

References

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

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

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

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

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

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

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

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

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

Oehrtman, M., Carlson, M. & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' functioning undersanding. In Carlson, M. P. & Rasmussen, C. (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 65-80), *MAA Notes #73.*

Parnafes, O. & diSessa, A. A. (in press). Microgenetic learning analysis: a methodology for studying knowledge in transition. *Human Development*.

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

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

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

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

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

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

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

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

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

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

COMPUTATIONAL THINKING IN LINEAR ALGEBRA

Spencer Bagley
University of California, San Diego

Jeffrey M. Rabin
University of California, San Diego

In this work, we examine students' ways of thinking when presented with a novel linear algebra problem. We have hypothesized that in order to succeed in linear algebra, students must employ and coordinate three modes of thinking, which we call computational, abstract, and geometric. This study examines the solution strategies that undergraduate honors linear algebra students employ to solve the problem, emphasizing the variety of productive and reflective ways in which the computational mode of thinking is used.

Key words: Justification, Linear Algebra, Problem Solving, Procedural Understanding

A first course in linear algebra plays a pivotal role in the mathematical education of college students in STEM disciplines. It usually follows a calculus sequence that may be predominantly computational in focus, and serves as a first encounter with many elements of more advanced mathematics. These include the emphasis on precise definitions and formal proofs, the creation of an abstract axiomatic structure, the study of objects that are not easy to visualize, multiple representations of those objects, and computational methods that require non-routine choices and interpretations of those representations. The kinds of flexible thinking required are foreign to many students, for whom "the fog rolls in" and the subject remains opaque (Carlson, 1993).

Much of the literature on the learning of linear algebra documents students' inabilities to solve basic problems, to move flexibly between the representations, to use abstract theorems in concrete situations, and even to speak or write the basic language of linear algebra coherently. Despite much insight into the causes of their difficulties, and creative pedagogical suggestions, the overall impression remains pessimistic.

Based on our own teaching experience, and consistent with previous work (Hillel, 2000; Sierpinska, 2000), we hypothesized that students must learn three major ways of thinking, and how to coordinate them with each other, in order to succeed in linear algebra. We term these *abstract*, *geometric*, and *computational* thinking. The study reported here was designed to explore students' use of these ways of thinking and their ability to coordinate them while solving a novel linear algebra problem. The initial research questions for the study were: What strategies do students use to solve a novel linear algebra problem (described below)? What are the uses, affordances, and constraints of each mode of thinking? In what ways do students coordinate these modes of thinking?

The students in our study were first year undergraduates enrolled in an Honors linear algebra course at a large university in the southwestern United States. As such, they do not form a representative sample of the general population of linear algebra students, and we would not expect our results to generalize immediately to this larger population. However, our students successfully used computational thinking to solve the problem in creative ways and to justify their solutions. Our study shows what successful student thinking in linear algebra can look like and provides a positive counterpoint to the literature on student deficiencies in linear algebra.

Theoretical Background and Literature Review

Drawing upon our own teaching experience, we hypothesize that mastery of linear algebra requires students to develop, and move flexibly between, three essential ways of thinking that we term abstract, geometric, and computational thinking.

Abstract thinking treats vectors as abstract objects manipulated according to formal rules, without reference to components or arrows. It makes use of definitions and theorems stated in coordinate-free language and applicable to $\mathbf{R}^n$ for any n. (Our course did not cover abstract vector spaces.) It makes assertions of existence or uniqueness on the basis of general principles rather than resulting from explicit computations. The notion of orthogonality may be part of abstract thinking if its meaning comes from an abstract inner product rather than the specifically geometric notion of right angles.

Geometric thinking involves visualization in Euclidean two- or three-dimensional space, with vectors represented as arrows, and may draw on concrete facts from high school geometry such as the Pythagorean Theorem. Span and linear independence can be geometric concepts if based on geometric properties such as collinearity and coplanarity in $\mathbf{R}^2$ and $\mathbf{R}^3$. "Orthogonal" as a synonym for "perpendicular" is a geometric concept. Systems of linear equations can be viewed geometrically in terms of the incidence relations among the lines or planes obtained by graphing each equation. A desirable outcome of a linear algebra course is for students to learn to extend their geometric thinking to apply in $\mathbf{R}^n$ for $n > 3$; we do not consider "geometric thinking" in such contexts to be an oxymoron.

Computational thinking represents vectors in $\mathbf{R}^n$ explicitly as n-tuples of real numbers, and draws conclusions from algorithms such as row reduction of matrices for solving systems or computing determinants. Assertions of existence or uniqueness come from explicit computations producing the objects in question. Systems of linear equations are thought of in terms of their coefficient or augmented matrices. Most importantly, for us computational thinking is not simply the procedural knowledge of how to execute an algorithm without errors. It includes choosing the appropriate computation to answer a particular question, understanding what the result of the computation means in that context, and reasoning about the steps or the outcome of the computation.

Problems in linear algebra may require an insight or reasoning process that is most accessible via one specific way of thinking. However, coordination between multiple ways of thinking is often required. By this we mean the ability to flexibly move from one way of thinking to another, or more specifically to "import" information acquired via one mode into another mode for further reasoning. Translation to the geometric mode often provides intuitive confirmation or understanding of abstract or computational results; a geometric picture may provide the key idea for an abstract proof or identify a key quantity to be computed.

Our tripartite taxonomy of modes of thinking is closely related to similar categories described by Hillel (2000) and Sierpinska (2000). Hillel speaks of three modes of description of the basic objects of study, namely vectors and operators, and calls these modes abstract, geometric, and algebraic. He traces many student difficulties to instructors' propensity for "constantly shifting modes of description and notation" without alerting students to this. He points out that choosing a basis, or changing basis, may require students to shift modes of description, and that over-reliance on a single mode can cause problems for students, for example when they inappropriately apply geometric intuition in dimensions higher than three. His three modes roughly correspond to ours, but note that they refer to *languages* that can describe the basic objects of study, and not directly to the ways that students think.

Sierpinska (2000), like us, discusses three modes of thinking or reasoning. We will shorten her names to (analytic-)structural, (synthetic-)geometric, and (analytic-)arithmetic, which roughly correspond in order to Hillel's categories. She says these modes are "equally useful, each in its own context, and especially when they are in interaction" (p. 233) and traces student difficulties to "their inability to move flexibly between the three modes" (p. 209). Her structural mode corresponds to our abstract one, and her geometric mode to ours. However, our computational mode seems to be more inclusive than her arithmetic one. She gives an example she considers to be intermediate between arithmetic and structural thinking, in which "the student was not performing calculations, he was reflecting on the properties of the possible effects of a calculation" (p. 240). We definitely consider this within the scope of computational thinking. Such *computational reasoning*, as we call it, will be a central theme of this paper.

We do not view computational thinking as merely procedural knowledge, in which students apply rote algorithms without reflection. We see it as an example of what Star (2005) calls *deep procedural knowledge*, which he characterizes as "knowledge of procedures that is associated with comprehension, flexibility, and critical judgment and that is distinct from (but possibly related to) knowledge of concepts" (p. 408). We will present examples of our students successfully applying computational thinking to solve the problem we posed in a variety of creative and reflective ways.

The literature on linear algebra tends to emphasize identification and diagnosis of student errors and deficiencies. Hillel (2000) and Sierpinska (2000) highlight their inability to utilize and flexibly coordinate the three modes of thinking. Dorier and Sierpinska (2001) also point out their unfamiliarity with axiomatic frameworks and the need to think in terms of formal definitions and general properties. Maracci (2008) studied student work on a challenging problem about the intersection of subspaces of a vector space of dimension at least 5. He suggested that their difficulties reflect an insufficiently general "paradigmatic model" of subspace limited to the span of a selected subset of given basis vectors, like the coordinate subspaces $x_i = 0$ of $\mathbf{R}^n$. He also pointed out the need to view a linear combination both as a process and as an object. Stewart and Thomas (2010) frame their work in terms of APOS theory and Tall's three worlds of mathematics (embodied, symbolic, and formal) which they compare to Hillel's three modes of description. They point out that students are often not given time and opportunities to develop links between the three worlds, and that their procedural knowledge is not deep in the above sense: "students who thought that they should row reduce a matrix often did not know why, or what to do with the result" (p.186).

Setting and Methods

The students in our study were enrolled in an Honors Linear Algebra course in their first year at a large university in the southwestern United States. The course forms the first quarter of a three-quarter Honors Calculus sequence in which linear algebra provides the conceptual basis for treating multivariable calculus in any number of dimensions. Completion of Advanced Placement calculus in high school, with the highest possible score of 5 on the AP calculus BC examination, is a prerequisite for the course. The instructor was not an author of this paper, although we did observe his class on a few occasions and asked him some questions about the course content. One of us (JMR) has taught the course in previous years. Eight students responded to our request for volunteers to participate in a clinical interview; all of these were accepted. Their course grades ranged from A though C. Although all volunteers were male, this was not unrepresentative of the class, which included only a few female students.

The interview revolved around the following task, termed the "Michelle Problem:"

> Michelle would like to create a basis for $\mathbf{R}^4$. She has already listed two vectors **v** and **w** that she would like to include in her basis, and wants to add more vectors to her list

until she obtains a basis. What instructions would you give her on how to accomplish this?

We had included this problem as one among many interview questions in an earlier pilot project (Wawro, Sweeney, & Rabin, 2011). Based on the responses at that time, it has the potential to elicit all three modes of thinking. Although some textbooks do prove that a linearly independent set of vectors can always be extended to a basis, this was not covered in the course our students took, and the problem was novel to them. The framing in $\mathbf{R}^4$ rather than $\mathbf{R}^3$ is intended to discourage an immediate appeal to visual geometric intuition, although geometric thinking is still applicable. Not providing numerical components for the "given" vectors may facilitate an abstract approach, and by asking for "instructions" for Michelle we hope that students will reflect on their methods and perhaps present them in algorithmic form.

Most of our students did not make much progress on the problem in this general setting. When they appeared "stuck" we asked them to test, or continue to develop, their ideas using the specific vectors $\mathbf{v} = [1\ 2\ 3\ 4]$ and $\mathbf{w} = [0\ \text{-}1\ 4\ 2]$. We asked a number of follow-up questions to probe students' intuition and solution procedures, whether students could formulate their solution procedure in an algorithmic way, and whether they could justify their procedure.

The interviews were videotaped and transcribed, and students' written work produced during the interviews was retained. These recordings, transcripts, and written documents formed the corpus of data analyzed in this study. Using grounded theory (Strauss & Corbin, 1994), we coded students' utterances as instances of abstract, computational, or geometric thinking, referring to written work for confirmatory evidence, and documented the ways in which students used each way of thinking.

Results

The students in our interviews were largely successful in solving this task; all but one of the students developed a solution procedure. The most frequently used solution method was some variant on the guess-and-check method; however, students usually provided strategically-informed guesses that they thought were more likely to succeed. A summary of solution methods and the number of students who used each is presented in the appendix.

Most of the students began with the abstract definition of basis before shifting into another mode of thinking. For instance, Carl's[1] initial response was, "She would just start by finding one vector that was linearly independent, that could not be made as a combination of **v** and **w**, and once she found that vector, she would have to find another one that was not a linear combination of the other three." Once they were given specific numerical vectors, students tended to quickly shift into computational thinking.

Students generally employed geometric thinking to provide intuition. Bob, for instance, initially employed a guess-and-check procedure, and mentioned that "the odds are fairly likely that they're not going to be linearly dependent." When pressed to explain this in more detail later in the interview, he reasoned geometrically by analogy with $\mathbf{R}^3$:

> *Bob*: She would just pick a third vector that's not in the same plane as the other ones. Which, if you're talking about comparing infinities, we're talking about a little infinity for just that plane, versus the huge infinity that's the rest of $\mathbf{R}^3$. So in $\mathbf{R}^4$ would be the same; there's a little three-dimensional space, and that the rest of $\mathbf{R}^4$ is huge, that you can choose from.

[1] All student names are pseudonyms.

Alan provides another example of geometric thinking powering intuition. When asked to describe orthogonality, he uses the abstract mode to provide a definition, then describes his intuition in more geometric terms:

> *Alan*: I usually think about it as being perpendicular, but I think the technical definition is more that the dot product is zero. But I like thinking about it as the directions in $\mathbf{R}^3$, just like your x and your y and your z pretty much. And then in $\mathbf{R}^n$, you just have n directions… they're sort of at right angles. I don't know how to think about that in $\mathbf{R}^n$, but that's what I think about; I just think about $\mathbf{R}^3$ and use that as a basis for my thinking about $\mathbf{R}^n$.

In contrast to the pilot study, where student thinking was largely balanced between all three modes of thinking, we noted a preponderance of computational thinking; indeed, one student provided no evidence of geometric thinking at all. We were thus led to ask, in addition to the initial research questions, in what productive ways students used computational thinking. The rest of this report focuses on some of the answers suggested by our data to this question.

Bob: Strategy generation and computational justification

To prompt students to justify their solutions, we included in our protocol the question, "Michelle is skeptical that your solution will always work. How would you convince her?" We had anticipated that this question would prompt abstract thinking, as this is the language commonly used in theorems and proofs. However, we found that several students were able to produce largely valid justifications using computational thinking alone. Additionally, several students' computational thinking inspired the creation of strategies for choosing vectors. The next student we will examine, Bob, is an example of both of these phenomena.

Bob used the Missing Pivots strategy, which he comes to formalize as follows: "Form an augmented 4x2 matrix with **v** and **w** and row reduce to echelon form. From there, choose two new vectors so that the nonpivotal rows have a nonzero number, and all other numbers in the vector are zero." In particular, Bob uses elementary basis vectors to supply the missing pivots.

This procedure was interesting to us in its own right for its sophistication and novelty; this was not one of the methods we had anticipated students would use. However, Bob's development of this procedure is also noteworthy. He first stated that "if you form a 4x4 matrix with [the vectors], it will row reduce to the identity in $\mathbf{R}^4$. So, you need to choose vectors that are likely to row reduce to pivotal columns." He then set out to find a way to produce such vectors, and eventually settled on the strategy described above. This strategy originated as a refinement of a guess-and-check method, or, in other words, a way to guess vectors that are more likely to succeed. His construction of this procedure could be described as reverse-engineering the check: by using computational reasoning to think about the test his vectors must pass, he was able to engineer vectors that are guaranteed to pass it.

Next, the interviewers asked him the "Michelle is skeptical" question. Bob's justification is entirely computational:

> *Bob*: Well, by the definition of linear independence, their matrix has to row-reduce to the identity. All the columns will be pivotal. So, by using these facts about linear independence and pivotal columns, this procedure is a way to find two more columns that will be pivotal columns independent of the other ones already found. With the two vectors she already has, she has two pivotal columns here, and they both represent pivotal ones. It's just a way to find the other two columns that won't form the same pivotal row as another one, so that they'll all be independent.

We note the presence of strongly computational language such as “row-reduce” and “procedure;” additionally, we argue that the notion of pivots on which this argument relies is so closely linked to the row reduction algorithm that it is a *de facto* sign of computational thinking.

This technique of justification by analysis of an algorithm is a particularly valuable way of constructing formal justifications. Many proofs in linear algebra typically proceed in a similar fashion; for instance, the usual proof that more than n vectors in $\mathbf{R}^n$ cannot be linearly independent proceeds by augmenting the vectors together in a matrix, row reducing, and arguing that the shape of the resulting echelon form implies that one vector is a linear combination of the others. The fact that students are capable of producing such justifications, as evidenced by Bob and others in our study, is perhaps an argument for teachers to highlight this method when discussing proof techniques.

It is worth noting that this justification falls slightly short of being a fully correct proof. In particular, say that the completed basis consists of the given vectors **v** and **w** and the additional vectors **x** and **y**. The vectors added to the row-reduced form of **v** and **w** are technically the row-reduced forms of **x** and **y**, and should have the inverse row operations applied to them to yield the original **x** and **y**. In Bob’s case, where the additional vectors are elementary basis vectors, the inverse row operations leave the additional vectors unchanged, but to be called a fully correct proof, his argument should have addressed this complication. Our data do not provide clear evidence of whether he was aware of this issue. In any event, the justification Bob produces is correct as far as it goes, and is most of the way to a complete proof.

Greg: Numerical examples and framing

The work of a student called Greg illustrates several more facets of students’ productive computational thinking. Here we will first discuss Greg’s solution procedure with minimal editorial comment, then offer two possible and related explanations of his actions.

Greg used the Unknown Columns strategy: he augmented **v** and **w** with vectors composed entirely of variables (see Figure 1), and performed row operations on the resulting 4x4 matrix until the first two columns were reduced to $\mathbf{e}_1$ and $\mathbf{e}_2$, obtaining the following matrix:

$$\left[\begin{array}{cc|cc} 1 & 0 & x_1 & y_1 \\ 2 & -1 & x_2 & y_2 \\ 3 & 4 & x_3 & y_3 \\ 4 & 2 & x_4 & y_4 \end{array}\right] \quad \begin{bmatrix} 1 & 0 & x_1 & y_1 \\ 0 & 1 & 2x_1 - x_2 & 2y_1 - y_2 \\ 0 & 0 & x_3 - 3x_1 - 4(x_2 + 2x_1) & y_3 - 3y_1 - 4(y_2 + 2y_1) \\ 0 & 0 & x_4 - 4x_1 - 2(2x_1 - x_2) & y_4 - 4y_1 - 2(2y_1 - y_2) \end{bmatrix}$$

Figure 1: Greg’s Unknown Columns matrix, initial and row-reduced

He then said, “These two [the diagonal entries] should be 1, and those [the off-diagonal entries] should be zeroes… because for these columns to be linearly independent, it would row-reduce to the identity matrix.” This reasoning gives him the following four equations for the third column **x**, and similar equations for the fourth column **y**:

$x_1 = 0$

$2x_1 - x_2 = 0$

$x_3 - 11x_1 + 4x_2 = 1$

$x_4 - 8x_1 + 2x_2 = 0$

The interviewers were interested to see how far he could push this line of reasoning, and asked him, “Can you give numerical vectors that satisfy these equations that you’ve got?” He obliged, and found that this system of equations uniquely determines the vector [0 0 1 0].

This appeared to violate his expectations; he asked, "Is there a way to take this back to the non-row-reduced form?" He thus seemed to believe that this was the row-reduced version of the original **x** vector, and was unsure how to recover the original vector. He also said he was sure that this row-reduced version of **x**, denoted **x**~, is independent of **v**~ and **w**~, but perhaps not independent of **v** and **w** themselves; further, he was sure that **x** (the original, "un-row-reduced" version) would be independent of **v** and **w**.

Later in the interview, the interviewers asked Greg to reflect on how he would validate a proposed third vector, and gave him a numerical example, [1 -1 0 0]. He augmented **v** and **w** with this vector and proceeded to row-reduce until his first two columns were $\mathbf{e}_1$ and $\mathbf{e}_2$. At this point, his third row was [0 0 | -15]. He seemed unsure how to interpret this result, saying, "I'm actually kinda confused about what this tells me. Um… Did I make a mistake?"

Once the interviewers assured him that he had not made a computational error, he reasoned further:

> *Greg*: Well, actually, if I continue row-reducing this, then I would get a 1 here, and then that would make it unsolvable. … And then, that would mean that this can't be a linear combination of these two. So then, it's not in the span.
> *Int*: And is that good or bad, for purposes of this problem?
> *G*: That's good. So then, this could be an additional vector for a basis.
> *Int*: And just what about your answer says that?
> *G*: If I continue row-reducing this, which is something that I also blanked out on, then I would get a 1 in the final column, and that states that 0 times whatever plus 0 times whatever, so on, equals 1, which is not true. So searching for solutions is what this is, so therefore there are no solutions.

Next, the interviewers asked him to reflect on what this example would mean for his original solution process (i.e., his Unknown Columns method). He realized that he had imposed too many constraints on the variables:

> *Greg*: I should have known that if this thing that ended up in the x_3 spot [i.e. the diagonal entry], if this was equal to 1, then there'd be no solution. So this is what I want to satisfy. … Really, that's the only important thing, basically. This stuff [i.e. the constraints on the non-diagonal entries] is not important. It's just this [constraint on the diagonal entry].
> *Int*: What do you mean it's just that that's important?
> <u>*G*</u>: This is the only thing that would determine whether or not my new column would be in the span. … And actually it's not just "equal to 1," it's "not equal to zero." So it can be any number.

Greg then replaced his previous conditions with the new condition:

$$x_3 - 3x_1 + 4x_2 - 8x_1 \neq 0$$

Figure 2: Greg's constraint

(It is evident in this picture that Greg first wrote "= 1" but then replaced this with "≠ 0".) He then chose numbers for the x_i's that satisfied the given constraint, and seemed much more satisfied with this solution than with his previous solution.

Now that we have recounted Greg's story, we propose two related explanations for his "aha moment" and subsequent correction of his strategy. First, the numerical example may have shown Greg that he was not done with his row reduction process. When row-reducing his matrix, he stopped as soon as the first two columns were in the form he wanted, and seemed to believe that the matrix was then in reduced row echelon form and should thus be equal to the identity matrix; note, for instance, that he obtains his initial equations by setting the third column of his matrix equal to [0 0 1 0], the third column of the identity matrix. He did not seem to realize that there was more work he needed to do to reach this point (i.e., dividing the third row by the diagonal entry to create a pivotal 1, etc.); this work may have been obscured by the notational complexity of his approach. In the numerical example, he obtained [1 3 -15 10] as his third column at the same point in the row reduction process, which seems to have helped him realize that the third column of his Unknown Columns matrix need not be equal to [0 0 1 0], allowing him to relax the constraints on the variables accordingly. It thus appears that numerical examples can illuminate what purely algebraic manipulations can obscure; however, both of these are examples of computational thinking. It is thus evident that computational thinking can take many forms, and that one form can be used to help make sense of another.

Additionally, the notational complexity of his work may have hindered Greg's thinking in another way. As shown by his desire to find the "non-row-reduced form" of his initial solution, he seemed to have lost track of the meaning of the variables he introduced; by definition, they are the entries in the *original* matrix, not its reduced form, and he had already undone the row reduction process by solving the equations he had accumulated in the third and fourth columns. Again, it appears that the numerical example helped him realize what his variables had originally meant.

Second, once Greg realized that he was not done, his thinking shifted from solving a system of equations (induced by setting his third column equal to [0 0 1 0]) to finding and satisfying constraints (induced by saying that the diagonal entry should be nonzero). This is an example of a shift in what we call a *framing*; that is, a student's mental expectation for how a computation should go. Greg shifted from thinking of *equations*, where the expected action is to *solve* and the expected outcome is a specific *solution*, to thinking of *constraints,* in this case inequalities, where the expected action is to *satisfy* and the expected outcome is a *range* of possible solutions. Both framings are useful at times, but for Greg, the framing as equation was less appropriate for the situation and seemed to cause him difficulty. The framing as equation led him to the (apparently unsatisfactory) conclusion that adding $\mathbf{e}_3$ and $\mathbf{e}_4$ should be the only solution to the problem; while this is a correct solution, it is not the most general form. By imposing unnecessarily strong constraints, he obtained a too-specific answer; by shifting his framing and relaxing the constraint, he obtained a much more general form.

Summary of affordances and pitfalls of computational thinking

Rather than comprehensively examine each student's uses of computational thinking, we here present as a more general summary a list of the affordances and pitfalls of computational thinking we identified in our students' thinking. We noted the following affordances of computational thinking:

- Working out an example to provide a general orientation to an unfamiliar problem: before being presented with specific vectors for **v** and **w**, Alan invented his own numerical examples to see how the other two vectors might be related
- Searching for a known algorithm that applies to the situation, or evaluating the applicability of a known algorithm: several students thought about using the Gram-Schmidt process, but eventually discarded this idea when they realized it required a different set of inputs than the situation afforded
- Recognizing when systems of equations have no solution, a unique solution, or infinitely many solutions: Greg (above) reasoned that his [0 0 15] row indicates that there is no solution, so the vector is not in the span of **v** and **w**; Hal said that

there are infinitely many bases that satisfy Michelle's requirements because there is not just one solution to a certain equation

- Clarifying a more general approach with a numerical example, as demonstrated above by Greg
- Making choices to simplify computations or reasoning: Bob proposed the strategic positioning of zeroes to "make it easier" to show that the guessed vectors are independent of the original vectors

Naturally, computational thinking was not a panacea; some students encountered difficulties that emerged directly from their computational thinking. Here are some of the pitfalls of this way of thinking that we observed:

- Thinking that coefficients in linear combinations must be integers
- Dividing by variable expressions without imposing the constraint that they must be nonzero
- Becoming overwhelmed by sheer algebraic complexity, e.g., many variables
- Circular substitutions leading to vacuous statements, such as 0 = 0
- Narrow focus on algorithmic procedures; for instance, some students thought that Gram-Schmidt was the only way to produce orthogonal vectors
- Uncertainty about meaning of variables, leading to difficulty in interpreting results
- Failing to realize when an algorithm has (not) been run to its proper completion
- Attaching significance to numerical coincidences

Many of these dangers will likely be familiar to any reader acquainted with students' thinking in linear algebra.

Discussion

Why did the students in this study use computational thinking so much more than the students in our pilot study? While our data provide no definitive answers, we can offer some conjectures. First, the class was taught with a heavy emphasis on computations. Many of the central problems of linear algebra were framed in terms of finding solutions to systems of equations; this is also the approach taken by the textbook (Hubbard & Hubbard, 2009). In contrast, the students in the pilot study were taught by a different instructor (JMR) who takes a more geometric approach to the subject.

Second, the structure of the interview may have privileged computational ways of thinking. To minimize feelings of frustration and to smooth the way for students to produce as much interesting mathematics as possible, we gave students the concrete numerical vectors fairly early on in the interview, which may have led students to more computational thinking than their natural inclination. We could well have waited until later to provide the concrete vectors, and pushed harder for non-numeric solutions to the problem.

Future work in this vein may attempt to remedy this shortcoming of the present study by retooling the protocol to more purposefully elicit multiple ways of reasoning. Several of our prompts appeared to privilege computational thinking, while others seemed to privilege geometric thinking. Further study in our own dataset will focus on identifying the prompts that reliably elicited specific types of thinking. Perhaps these prompts could be purposefully modified and combined to provide a more complete picture of the ways students are able to use all three modes of thinking.

We may recommend a few pedagogical implications of our work. First, the success of students in constructing computational justifications by analysis of algorithm (and the prevalence of canonical proofs that proceed in this manner) may imply that instructors should foreground this technique in classes with a proof component. Second, our results suggest that instructors may wish to adopt a more positive orientation toward computational and procedural thinking in general, notice when their students are using it productively, and help them grow more expert-like in the ways they use it.

Finally, we wish to align ourselves with the work of Star (2005) in rehabilitating the notion of procedural knowledge. We recognize that because it is composed of students in a high-achieving population, our sample is not representative of the general population of linear algebra students, but we regard our results as an existence proof. Our students' use of procedural elements such as framing and reverse-engineering in the service of activities such as strategy generation, justification, and troubleshooting is evidence that linear algebra students can, perhaps with appropriate scaffolding from their instructors, use procedural knowledge in mathematically sophisticated ways. We hope that other researchers and practitioners will join us in recognizing and researching productive ways students can reason procedurally.

References

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

Dorier, J.-L., & Sierpinska, A. (2001). Research into the teaching and learning of linear algebra. In Holton, D. (Ed.), *The Teaching and Learning of Mathematics at University Level: an ICMI Study* (pp. 255-273). Dordrecht: Kluwer Academic.

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.

Hubbard, J., & Hubbard, B. B. (2009). *Vector calculus, linear algebra, and differential forms: A unified approach.* Ithaca, NY: Matrix Editions.

Maracci, M. (2008). Combining different theoretical perspectives for analyzing students' difficulties in vector spaces theory. *ZDM Mathematics Education, 40*, 265-276.

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.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*(5), 404-411.

Stewart, S., & Thomas, M.O.J. (2010). Student learning of basis, span and linear independence in linear algebra. *International Journal of Mathematical Education in Science and Technology*, 41(2), 173-188.

Strauss, A., & Corbin, J. (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.

Wawro, M., Sweeney, G., & Rabin, J. (2011). Subspace in linear algebra: investigating students' concept images and interactions with the formal definition. *Educational Studies in Mathematics, 78(1),* 1-19.

Appendix: Solution Methods for the Michelle Problem

When we designed our study, we listed possible solution methods that we expected students might use. These are named in **boldface** below, and are followed by the additional unanticipated methods our students used. The numbers of students observed to employ each method is given. Note that some students used multiple approaches.

1. **Guess and Check** (5 students). Simply guess two vectors **x** and **y** to supplement **v** and **w** in forming a basis. Check that the set of four do form a basis, for example by row reducing a 4x4 matrix having these vectors as columns. This method succeeds with probability one, that is, unless the guesses are very unlucky.

2. **Abstract Proof** (0 students). Since the span of **v** and **w** is two-dimensional, choose any vector **x** not in this span. Then choose any vector **y** not in the three-dimensional span of **v**, **w**, **x**. This is a proof of existence of the desired basis, in the Abstract language. To translate it into a practical algorithm, Computational thinking is needed to specify a method for making the required choices. For example, the span of **v** and **w** can be characterized by row reducing a 4x2 matrix having those columns, to find a simple basis for this span. Then find by inspection a vector that is not a linear combination of these basis vectors. Or, augment this matrix with a third column of unknowns and row reduce to determine the conditions on the unknowns for this third column not to be a linear combination of the first two. Several students stated the Abstract definition of a basis, but did not appear to extend this into a proof by, e.g., arguing that there are vectors in $\mathbf{R}^4$ outside the two-dimensional span of **v** and **w**.
3. **Orthogonality** (0 students). Find all vectors orthogonal to both **v** and **w** (Geometric thinking), for example by computing the kernel of the 2x4 matrix having those rows. Choose two independent vectors in this kernel to complete the desired basis. This method was suggested to students during the interview, but none used it spontaneously. This was in contrast to the pilot study, in which many students proposed it as their initial response.
4. Modified Linear Combination (1 student). Take some linear combination of **v** and **w**, and alter one component of the resulting vector. This gives a vector **x** that is (almost certainly) independent of **v** and **w**. Repeat with some linear combination of **v**, **w**, and **x**.
5. Unknown Columns (4 students). Create a 4x4 matrix whose columns are **v**, **w**, and two columns of unknowns representing the desired additional vectors **x** and **y**. Row reduce to determine the conditions on the unknowns that make the four columns independent. Choose values satisfying these conditions.
6. Missing Pivots (2 students). Row reduce the 4x2 matrix having columns **v** and **w**, noting which rows of the reduced matrix contain pivots. Then supply two additional columns having pivots in the complementary rows (for example, two of the standard basis vectors would accomplish this). The reduced columns are then independent, and by reversing the sequence of steps in the row reduction one obtains four independent columns **v**, **w**, **x**, **y** forming a basis. The choice of standard basis vectors is particularly convenient because they will often not change under the reduction steps.

Understanding Mathematical Conjecturing

Jason K. Belnap and Amy L. Parrott
University of Wisconsin, Oshkosh

In this study, we open up discussions regarding one of the unexplored aspects of mathematical sophistication, the inductive work of conjecturing. We consider the following questions: What does conjecturing entail? How do the conjectures of experts and novices differ? What characteristics, behaviors, practices, and viewpoints distinguish novice from expert conjecturers? and What activities enable individuals to make conjectures? To answer these questions, we conducted a qualitative research study of eight participants at various levels of mathematical maturity. Answers to our research questions will begin to provide an understanding about what helps students develop the ability to make mathematical conjectures and what characteristics of tasks and topics may effectively elicit such behaviors, informing curriculum development, assessment, and instruction.

Key words: Mathematical sophistication, Enculturation, Mathematical behaviors, Conjecturing

Background

Much of contemporary research is concerned with helping students become more adept at problem solving and learning mathematical ideas. In fact, an ongoing concern is empowering students to develop deep understanding of mathematical concepts, instead of only developing shallow procedural proficiency (Carpenter & Lehrer, 1999; Davis, 1992; Kazemi & Stipek, 2001; Lampert, 2001; Lockhart, 2008; Schoenfeld, 1988); that is, we want students to be able to apply their knowledge to solve new problems (Devlin, 2000; Halmos, 1968).

One key to addressing these challenges lies in understanding the practices of the mathematical community. From a sociocultural perspective (Bauersfeld, 1995; Cobb & Yackel, 1996), individuals develop their own mathematical understanding as they contribute to their local mathematical culture and engage in the practices of the larger mathematical community. Citing Cobb, Bowers, Lave, and Wenger, Rasmussen, Zandieh, King, and Teppo (2005) argue that learning mathematics is synonymous with participation in mathematical practices; in other words, many of the activities used by the mathematical profession to build new mathematical artifacts are needed by learners to acquire those same artifacts. Carlson and Bloom (2005) argued that successful problem solving involves more than content knowledge; it requires cognitive control skills, methods, and heuristics.

Schoenfeld (1992) noted that novices lacked the skills and behaviors characteristic of expert mathematicians, skills which go beyond simple content knowledge, such as: attending carefully to language, building models and examples, making and testing conjectures, and making arguments based on the structure of a problem. Seaman and Szydlik (2007) observed this when they observed that preservice teachers failed to relearn forgotten, common, elementary school mathematics concepts and skills (even when given ample time and resources) because they lacked *Mathematical Sophistication* (the habits of mind and practices of the mathematical community); they noted that mathematical sophistication includes: making sense of definitions, seeking to understand patterns and structure, making analogies, making and testing conjectures, creating mental and physical models (examples and non-examples of things), and seeking to understand why relationships make sense. They also found evidence that basic levels of mathematical sophistication can be measured and can be developed during the course of a class (Szydlik, Kuennen, Beam, Belnap, & Parrott, 2011).

Mathematical sophistication is thus not only critical for prospective mathematicians, but for *anyone* who must engage in mathematical learning and problem solving. It is these mathematically sophisticated behaviors (such as conjecturing, testing, and modeling) that empower problem solvers to correct their own models and arrive at solutions (Moore, Carlson, & Oehrtman, 2009). Enculturation into the practices of the mathematical community becomes a key to becoming an effective mathematics learner. So, in order to empower students to more effectively learn mathematics, we must understand the practices of the mathematical community and find ways for them to acquire these practices. As we identify those practices that enable mathematicians to *do* mathematics and find ways to instill these in our students, we will empower them mathematically.

Professional mathematicians expand the community's mathematical knowledge by engaging in a variety of mathematical activities, most of which could be classified as inductive or deductive work. The inductive work of mathematics involves activities that generate new mathematical ideas; through investigation, exploration, or study, mathematicians create conjectures (i.e. new and unproven hypotheses). Through deductive work, mathematicians take assumed mathematical ideas (axioms) and proven mathematical facts (theorems) and either prove or disprove those conjectures. Proven theorems then become part of the community's mathematical knowledge (i.e. becoming new mathematical artifacts).

In this study, we have chosen to focus our efforts on the mathematical activity of conjecturing for three main reasons. First, conjecturing is critical to the field of mathematics; it represents the potential mathematical knowledge of the field. Second, conjecturing tends to be a neglected aspect of mathematics classes and enculturation. Most lower-level mathematics courses focus on understanding and applying known theorems. At the upper-level, most attention is devoted to proofs and logic because of their complex and problematic nature, thus focusing predominantly on understanding content, testing conjectures, developing logical arguments, developing counterexamples, and writing deductive arguments--a focus mirrored by the research literature (Weber, 2001; Selden & Selden, 2003; Alcock & Weber, 2005). Third, (in theory) conjecturing is more accessible to undergraduates and novices because it is the generation of ideas which need not be certain; it is hypothesis-making and does not require all the intricacies that accompany deductive work.

This paper focuses on addressing the following question: What does mathematical sophistication look like for conjecturing? To answer this question, we consider the following questions: What does conjecturing entail? How do the conjectures of experts and novices differ? What characteristics, behaviors, practices, and viewpoints distinguish novice from expert conjecturers? and What activities enable individuals to make conjectures?

Methodology

As an entry point to understanding conjecturing, we chose to examine individual conjecturing practices outside of the complexity of the classroom setting. We conducted a qualitative research study, during spring semester 2012, in a mathematics department at a public university. In order to contrast the individual conjecturing practices of a diverse set of individuals, we purposively selected eight participants at various levels of mathematical maturity: two undergraduate students (novice mathematicians), Scott and Laura; three graduate students (apprentice mathematicians), Ann, Noah, and Charlie; and three research mathematicians (i.e. expert mathematicians) Sam, Josh, and Lisa. We chose student participants from a list of volunteers, who were enrolled in one of two specific courses; undergraduate students were taking a 200-level introduction to proofs course, while graduate students were enrolled in a graduate course in advanced Euclidean geometry, which incorporated a conjecturing component.

As shown in table 1, our participants varied by gender, geometry background, mathematical ability, and position in the mathematical culture. Our participants represented different levels of enculturation into the mathematical community, namely: two undergraduate students (i.e. novices), three graduate students (i.e. apprentices), and three research mathematicians (i.e experts). We purposively selected students for diversity in ability (high/medium/low) to do authentic mathematical work, as judged by their instructors and (only in the case of graduate students) conjectures previously produced on conjecturing tasks. Mathematicians were selected to represent diverse areas of mathematical expertise, which included: probability theory, graph theory, and social choice theory.

Table 1. Study Participants

Participant		Level/Expertise	Task Time	GeoGebra Use
[a]Novice (UGrad)				
Scott	M	Low	57 min	Interview only
Laura	F	Middle	1 hr 28 min	Yes
Apprentice (Grad)				
Ann	F	Low	1 hr 26 min	Yes
Noah	M	Middle	1 hr 18 min	Yes
Charlie	M	High	50 min	Yes
Expert (Research)				
Sam	M	Social Choice Theory	1 hr 21 min	Yes
Josh	M	Probability Theory	1 hr 58 min[b]	Yes
Lisa	F	Graph Theory	1 hr 16 min[b]	Yes

Note: [a]No highly ranked novices were willing to participate in the study. [b]Josh and Lisa were stopped by the interviewer(s) for time's sake.

Data Collection

Data collection revolved around each participant's involvement in the individual conjecturing task shown in figure 1. Participants were presented with a hardcopy of the task and given ample time and resources to work on it, then (after a break) participated in an interview regarding their experience, approach, and conjectures.

The task was one of a set of nine conjecturing tasks that we had created to add a conjecturing component to the aforementioned graduate geometry course. We chose to use this particular task for this study because it involved an exploration of three non-conventional geometric objects, presenting a novel exploration for experts and apprentices, as well as novices.

Figure 1: Conjecturing task completed individually by each participant.

Consider a generic quadrilateral $ABCD$ and the three types of derived quadrilaterals: the angle bisector quadrilateral, the midpoint quadrilateral, and the perpendicular bisector quadrilateral. Your task is to explore the relationships between quadrilateral $ABCD$ and each of these different types of derived quadrilaterals; make conjectures based upon your observations.

Task 3: Formal Definitions of Derived Quadrilaterals

Angle Bisector Quadrilateral: The *angle bisector quadrilateral* of a quadrilateral $ABCD$ is a quadrilateral $A'B'C'D'$, where A', B', C', and D' are each the intersection of the angle bisectors of the corresponding and subsequent vertices (e.g. A' is the intersection of the bisectors of $\angle DAB$ and $\angle ABC$).

Midpoint Quadrilateral: The *midpoint quadrilateral* of a quadrilateral $ABCD$ is a quadrilateral $A'B'C'D'$, where A', B', C', and D' are the midpoints of $\overline{AB}, \overline{BC}, \overline{CD}$, and $\overline{DA}$ (respectively).

Perpendicular Bisector Quadrilateral: The *perpendicular bisector quadrilateral* of a quadrilateral $ABCD$ is a quadrilateral $A'B'C'D'$, where A', B', C', and D' are each the intersection points of the perpendicular bisectors of the two adjacent sides (e.g. A' is the intersection of the perpendicular bisectors of $\overline{DA}$ and $\overline{AB}$).

Note: None of these are conventionally defined in the mathematics community and of these, only the *midpoint quadrilateral* is commonly encountered. Only the *midpoint quadrilateral* is defined for all quadrilaterals; the others have undefined (or degenerate) cases.

The task was intended to create a context where participants could do inductive work leading to the generation of conjectures. One faculty participant had previously encountered the midpoint quadrilateral, otherwise all three definitions were novel for each participant. Participants were allowed to explore these definitions, with the goal of writing as many conjectures about them as possible; they were given as much time as they desired, with the exception of Josh, whom we cut off after about two hours, and Lisa, whom we cut off after approximately 75 minutes.

During the task, participants had open access to the following resources: a ruler and compass; various colored and regular writing instruments; paper; a list of Euclidean postulates, common notions, definitions, and propositions; a glossary of common geometry terms; and GeoGebra, a free dynamic software program for constructing, measuring, and manipulating dynamic geometric objects. Researchers were also available throughout to answer questions on the task and definitions or to help with software usage.

After the task, each participant took a break, during which we prepared for the interview by discussing our observations and the participant's conjectures and making adjustments to the interview questions. The subsequent interview focused on: a) clarifying any behaviors and thinking that were not discernible by outward observation, b) understanding the participant's experience and perspective, c) understanding the written conjectures, and d) gaining details about the conjecturing process.

Data consisted of video recordings, written work, and observation notes. We took three video recordings during each task and subsequent interview: a) a video taken from the side, of them working on the task; b) a top-down recording of their written work; and c) an internal recording of their computer work. A synchronized compilation of these three video recordings served as the primary data source. Secondary data sources included each participant's written work and conjectures along with our observation notes taken during each task and interview.

Data Analysis

We analyzed the data using grounded theory techniques (Strauss & Corbin, 1998). Beginning with Scott (low-level novice), we independently reviewed the coordinated video feeds and made time-stamped annotations describing his behaviors throughout the task. Then we compared our annotations and negotiated differences, using the video and secondary sources of data to triangulate our observations. We began clustering our annotations around common themes to form initial categories of behaviors. We utilized our results to inform our next interview and observations (of Laura).

We repeated this same process working upward (by experience and level) through our participants by next meeting with Laura (medium-level novice) and then Ann (low-level apprentice). After meeting with each participant, we coded independently by writing time stamped annotations and applying our emerging categories; doing so, we modified our categories as appropriate to accurately describe the data; each time we made changes, we back-coded prior participant data to examine how the framework reflected the data.

By our meeting with Noah (medium-apprentice), we noticed some broad, over-arching themes, so we independently synthesized our analyses of Noah into a synopsis (vignette) by describing the characteristics and dimensions of each theme. We continued this process, independently writing, then collaboratively negotiating a final synopsis for each participant. As we compiled these synopses, we did a cross-case analysis, examining differences and similarities across the different levels of expertise. In this paper, we present some of these differences as dimensions of the mathematical activity *conjecturing* and illustrate these by contrasting two cases: Scott (novice) and Josh (expert).

Results

Five themes emerged from our data analysis, which describe ways in which participants differed in their conjecturing practice. First, participants differed in the ***overall conjecturing process*** that they employed, that is how they approached the task, and what strategies they employed in order to go from the task to a set of formal conjectures. Second, there was a diversity among participants in the ***objects*** that they ***created for investigation***, that is what types of models or examples they created to aid their investigation and ***how*** they viewed or used them. Third, the ***nature of observations*** made by participants varied; this includes what they paid attention to and what they chose to observe. Fourth, there were differences in the ***qualities of written conjectures*** created by the participants; these differences extend beyond the properties stated to encompass differences in mathematical quality of the statements themselves, including semantic and syntactic qualities. Finally, there was a difference in the ***qualifications of written conjectures***, that is differences in what the individual considered to be worthy of being considered a conjecture or at what point they were willing to commit an idea to paper as a formal conjecture.

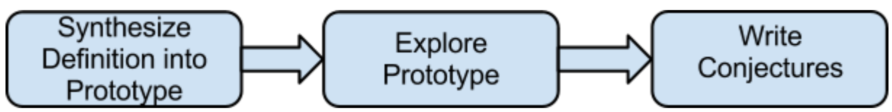
In this paper, we contrast two extreme cases, which illustrate some of these differences. The first case is Scott, our most novice conjecturer; and the second case is Josh, an expert researcher in the field.

Case of a Novice Conjecturer -- Scott

Scott is a male undergraduate student, majoring in both mathematics and business. His last experience with geometry (in high school) occurred approximately five years before our study. His instructor ranked him in the lowest third for his ability to do authentic mathematical activities. According to his background questionnaire, he had never taken a course requiring him to explore mathematical ideas and then make conjectures.

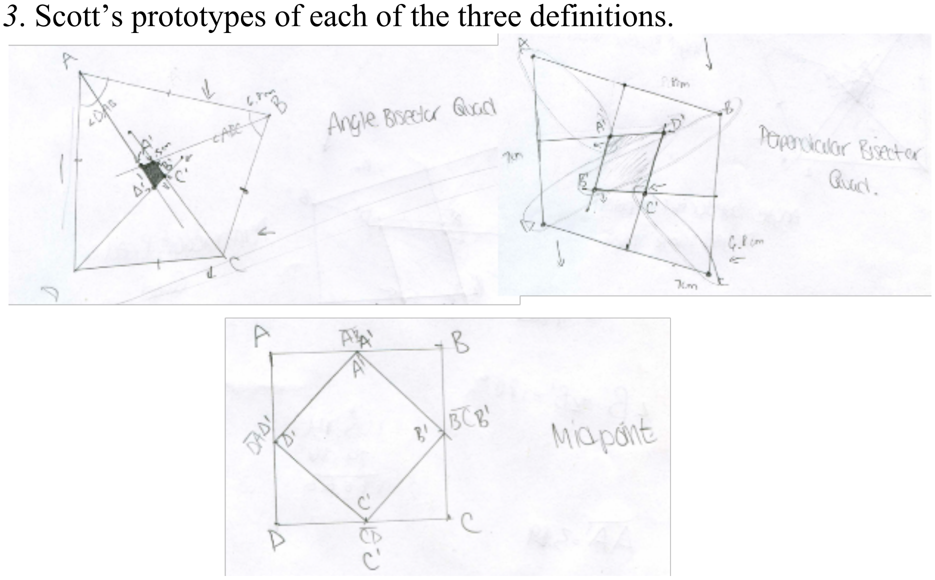
Scott had a very consistent and methodical approach (shown in figure 2) to the conjecturing task, which he applied to each definition. Scott began by creating a ***prototype***, synthesizing the definition into a precisely constructed drawing. He did so by drawing a quadrilateral and (while reading the definition) carefully drawing the quadrilateral derived from the definition, ensuring appropriate angle and length measurements by using a ruler and a protractor.

Figure 2. Scott's approach to the conjecturing task.



Scott was unable to completely synthesize the definitions of the angle and perpendicular bisector quadrilaterals without assistance, because he did not know the meaning of ***angle bisector*** and ***bisector***. We intervened, explaining what an angle bisector means and verifying his attempts to construct the first vertex of his angle bisector quadrilateral prototype. After that, he understood enough to complete this prototype, and the subsequent ones (shown in figure 3).

Figure 3. Scott's prototypes of each of the three definitions.

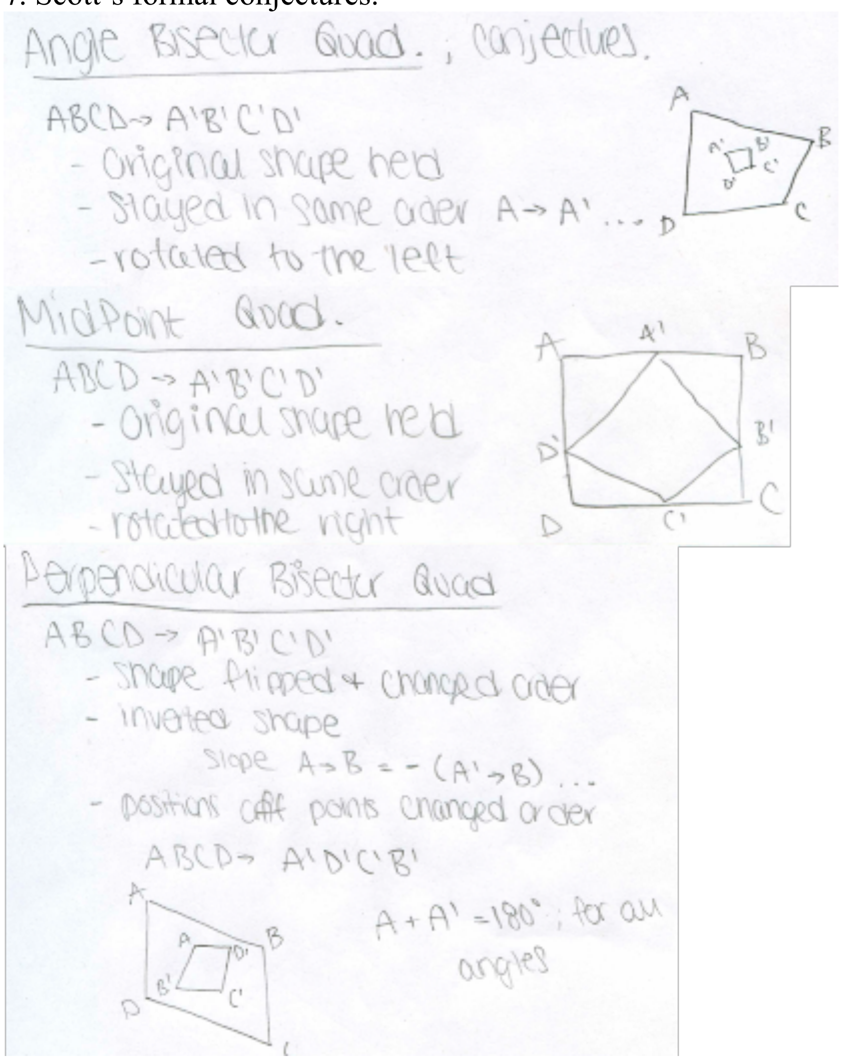


Note: Each of the three prototypes was drawn on a separate page. Congruency marks and circled pairs of vertices were added during exploration (task, 27 Mar 2012).

After creating his prototypes, he examined them. Scott's examination process involved switching between prototypes, physically rotating the figures, comparing the quadrilaterals and derived quadrilaterals, measuring and comparing side lengths, and measuring the distance between corresponding vertices (synopsis, 2 Jul 2012).

Once he finished examining the prototypes, he wrote down all of his conjectures at once, which are shown in figure 4. In his conjecturing work, Scott focused on superficial features of his prototypes, such as orientation on the page, how the shape appears (i.e. what it *looks like*), and the orientation of the vertices (i.e. clockwise or counter-clockwise). His conjectures were incomplete statements, often using imprecise language and even the unconventional use of some mathematical terms (e.g. rotated and inverted.)

Figure 4. Scott's formal conjectures.



Note: The final conjecture about the relationship between the corresponding angles of a quadrilateral and its perpendicular bisector quadrilateral was added during the debriefing interview and was not part of his work on the conjecturing task.

In our debriefing interview (27 Mar 2012), Scott stated that he felt the task was difficult for him to complete. He considered the task difficult because there was no specific targeted end result. He was unsure of what he should be observing, what features should stand out in the shapes, and what he should look for or check.

Case of an Expert Conjecturer -- Josh

Josh is a male research mathematician and faculty member specializing in probability theory. He has taught courses in Euclidean and non-Euclidean geometries to prospective elementary and middle level teachers.

As illustrated in figure 5, Josh exhibited a very non-linear approach to the task, alternating between a variety of problem-solving strategies. Like Scott, Josh's entry point into the task was synthesizing the definition into an example, but he did so in two different ways. One was to create an initial sketch of one example of the object. The other was to go directly to the construction of a dynamic model. Once an initial model was constructed, he used metacognitive abilities to shift among different practices, which included: constructing and manipulating dynamic models, pausing for prolonged periods of time to review his creations or take a break, sketching special cases, writing down questions and informal conjectures, deductively reasoning with known facts, and writing down formal conjectures.

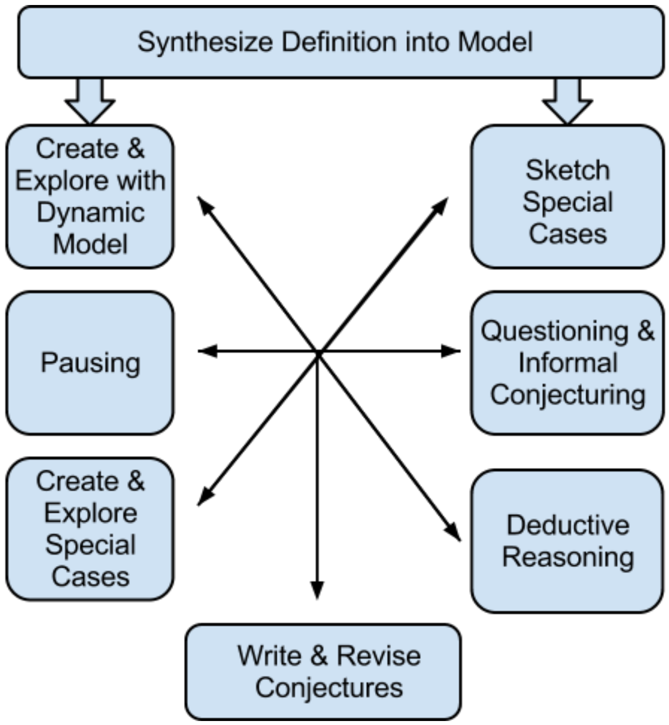
Throughout his process, Josh was strategic. He was imprecise in his drawings, treating them as figurative representations of ideas he had in his mind. He was also strategic and intentional in his work. He intentionally examined special cases, such as conventional classes of quadrilaterals. He also spent a great deal of time and effort looking for properties and cases where the definitions would be undefined. He frequently wrote down questions or informal conjectures and then hunted for examples and situations that addressed them.

Josh's overall process was dynamic and constructive. He frequently changed directions, shifting between various activities. He also re-engaged ideas and examples he had previously written. Even his formal conjectures were subject to exploration and scrutiny; he revisited and revised them as he encountered new evidence. At the end of the task, Josh still had ideas for exploration; he stated that at the beginning of the task, his first thought was to examine if the three definitions were related, but he ran out of time to examine this possibility (interview, 14 Jun 2012).

In the interview (14 Jun 2012), Josh claimed that the process we observed was not as typical of his professional work. When he does his research, he will find an interesting ideas and explore it further because he wants to pursue it, not specifically to find conjectures; he feels that this, "makes [him] a better conjecturer (interview, 14 Jun 2012)." Also, in his work, he has good ways of keeping track of his work in his head, so he does more pausing to think and explore in his head and writes less down.

Josh's conjectures were characterized as refined, mathematically precise statements. He also had an intuitive standard for something to be considered a conjecture, that it should be interesting and not obvious; it needs to be "worth writing down" (interview, 14 Jun 2012). He even stated during our interview that one of his conjectures should be crossed out because it was trivial. Furthermore, Josh needed to see some avenue of proof before he would write down an idea as a formal conjecture.

Figure 5: Josh's approach to the conjecturing task.



Discussion

Our study reveals conjecturing as a highly complex and holistic activity, entailing a variety of skills, knowledge, and values. It is affected by logic, content knowledge, practices of observation, experience with mathematical language, and persistence. Indeed, conjecturing appears to rely upon many of the characteristics Seaman and Szydlik (2007) and Schoenfeld (1992) identified.

Josh's case illustrates this complexity, as he drew upon various aspects of his cognition and mathematical practice. Josh was empowered to generate many conjectures because he created and considered a wide variety of examples. He utilized his understanding of conventionally defined geometric objects and properties both to subdivide the task into special cases for consideration and to express the conjectures and properties that he observed. He drew upon his mathematical experience to scrutinize the definitions and hunt for problematic examples and degenerate cases. He used "expert executive skills" (Schoenfeld, 1992) and metacognition to coordinate a variety of problem solving strategies. He utilized both inductive and deductive approaches to hunt down relationships and properties of the objects involved. He also used his experience and understanding of mathematical language to articulate precise mathematical conjectures.

By contrast, Scott was limited in his conjecturing (both in what he could create and articulate) by his limited skills, understanding, and mathematical practices. He was unable to generate examples of defined objects because he lacked an understanding conventional vocabulary. Once the vocabulary problem was remedied, he was further limited by his prototypical view, literal treatment of examples, and limited mathematical language skills.

Yet in spite of his limitations, Scott (a novice) was able to generate nine conjectures, most of which were not (but potentially could be, with assistance) formulated into precise, testable statements. Furthermore, with the intervention of dynamic geometric software, he was able to create a more sophisticated, tenth conjecture, which he stated more precisely than his other conjectures.

We conjecture that *conjecturing* could provide an important context for both learning mathematics and enculturation. In spite of its potential complexity, Scott's case suggests that conjecturing may be accessible to novices; differences in individuals' mathematical sophistication and experience may simply shape or limit the type of attainable conjectures. Its accessibility to novices suggests that it could be a powerful experience for individuals located at the periphery of the profession to gain authentic mathematical experience. Because of the holistic nature of conjecturing, it may provide a context for establishing important sociomathematical norms and corresponding individual mathematical beliefs regarding such things as: the role and use of mathematical definitions, language usage (e.g. conjunctions, conditional statements, and quantifiers), example use and selection, and distinctions between evidence and proof.

Based on our socioconstructivist position, the possibility of implementing conjecturing into mathematics coursework to facilitate mathematical development and enculturation raises at least three important topics for further inquiry. A first topic of inquiry involves classroom culture. Participants in our study worked individually on a conjecturing task; however, individual involvement in the classroom is closely linked to classroom norms and practices, raising such questions as: What does productive conjecturing look like in the classroom? How do (or can) individuals *collaboratively* approach conjecturing? What norms and practices are involved in conjecturing? and What norms and practices foster conjecturing?

There are several ways that studies could approach such research questions. An initial investigation of such questions might involve contrasting the approaches groups of students and groups of experts make when collaboratively engaging in a conjecturing task, then examining how the norms and practices of these two groups differ. From there, a set of ideal norms might be projected for the classroom and a design experiment could be conducted to understand how such norms could be fostered and developed through various experiences, activities, and tasks. Such a design experiment might additionally (or alternatively) focus on understanding which mathematical practices and norms can be established in the context of conjecturing work and how that might be done. Finally, classroom studies could utilize individual (pre and post) conjecturing tasks in order to identify changes in the conjecturing behaviors of participants due to implemented treatments.

Building on classroom culture, a second topic of inquiry is the task itself. Reflecting on our own experiences and contrasting those with our participants, we have reason to believe that the nature of the task strongly impacts conjecturing behaviors. We originally designed the task (along with eight others) to infuse a conjecturing component into an advanced graduate course in Euclidean geometry. As we did so, we thoroughly worked each task individually in order to generate a list of potential conjectures, and to determine each one's validity for inclusion in a set of instructor materials. As we reflected back on our own practices, we noticed that *we* often incorporated both inductive *and* deductive reasoning to generate our conjectures. We worked to accumulate knowledge about the objects and even made related and connected conjectures about each.

In the present study, we expected experts to take a similar approach, but while this did happen in a few instances, we were surprised that it did not happen more. Josh exhibited these behaviors twice. Once, he made a conjecture that built upon another conjecture. Another time, he used deductive reasoning on one of his drawings to pose a conjecture regarding the relationship between the corresponding angles of a quadrilateral and its perpendicular bisector quadrilateral.

Although it is possible that participants *were* cognitively attempting to accumulate understanding of the objects, we believe that the task did not foster connected thinking. The task puts the stating of conjectures as the direct end result of the activity. We conjecture that individuals' activities would be different if the accumulation of knowledge or understanding was the task's goal. For instance, a task which places individuals in the hypothetical situation of preparing to teach a class about these objects would likely push participants to generate connected knowledge, although it would also diminish conjecturing in favor of notions of proof.

This notion of task design raises important questions for inquiry. These include: What types, characteristics, or features of tasks make them well suited for conjecturing? How do differences in tasks relate to differences in individual (or collective) conjecturing practices? What tasks might provide more authentic conjecturing experiences? Such questions could be treated as a topic for the design studies previously suggested.

Finally, a third line of inquiry, closely related to task development and classroom implementation, is mathematical domains. In this study, our task involved studying objects defined in the context of Euclidean geometry; these objects were easily constructed, observed, and manipulated (especially with the available software). Because most of the resulting conjectures described visually apparent or observable relationships and properties, we conjecture that much of the task's accessibility may be due to the contextual domain. We wonder what such tasks would look like in other domains (e.g. group or ring theory, topology, or analysis) and how accessible they would be, particularly when involving objects that are abstract or not easily represented graphically. This change in the nature of objects (as well as available tools) would plausibly require a different conjecturing approach. A study structured similar to this one, but with a different contextual domain could provide valuable insights when contrasted with our results and would be a likely starting point for such inquiry.

Due to its generative nature, conjecturing is a critical mathematical practice. It is important to the growth of the field that individuals develop this practice. In this report, we demonstrated that while conjecturing at the expert level may entail great complexity, it can be accessible for even the most novice mathematician. Hence, conjecturing provides an avenue for participation in authentic mathematical explorations. While further research is needed to determine specific characteristics of tasks or behaviors that will aid in the development of the ability to make conjectures, we believe we have begun a crucial step in the process of understanding conjecturing.

References

Alcock, L, & Weber, K. (2005) Proof validation in real analysis: Inferring and checking warrants. *The Journal of Mathematical Behavior*, *24*, 125-134.

Bauersfeld, H. (1995). A constructivist approach to teaching. In L. Steffe & J. Gale (Eds.), *Constructivism in Education* (pp. 137-158). Hillsdale, NJ: Erlbaum.

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58 (1)*, 45- 75.

Carpenter, T., & Lehrer, R. (1999). Teaching and Learning Mathematics With Understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*, 175–190.

Davis, R. B. (1992). Understanding "Understanding". *Journal of Mathematical Behavior, 11,* 225-241.

Devlin, K. (2000). *The Language of Mathematics: Making the invisible visible*. New York: W. H. Freeman and Company.

Halmos, P. (1968). Mathematics as a creative art. *The American Scientist, 56*, 375-389

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

Lampert, M. (2001) Teaching While Leading a Whole-Class Discussion. In *Teaching Problems and the Problems in Teaching* (pp.143-177). New Haven, CT: Yale University Press.

Lockhart, P. (2002). A mathematician's lament. Retrieved from http://www.maa.org/devlin/LockhartsLament.pdf

Moore, K. C., Carlson, M. P., & Oehrtman, M. (2009). *The role of quantitative reasoning in solving applied precalculus problems.* Paper presented at the Twelfth Annual Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education (SIGMAA on RUME) Conference, Raleigh, NC:North Carolina State University.

Rasmussen, C., Zandieh, M., King, K., & Teppo, A. (2005). Advancing mathematical activity: A view of advanced mathematical thinking. *Mathematical Thinking and Learning*, *7*, 51-73.

Schoenfeld, A. H. (1988). When good teaching leads to bad results: The disasters of well taught mathematics classes. *Educational Psychologist, 23*(2), 145-166.

Schoenfeld, A.H. (1992). Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 334-370). New York: Macmillan Publishing Company.

Seaman, C., & Szydlik, J. (2007). Mathematical sophistication among preservice elementary teachers. *Journal of Mathematics Teacher Education, 10*, p. 167-182.

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.

Strauss, A., & Corbin, J. (Eds). (1998). *Basics of Qualitative Research: Techniques and procedures for developing grounded theory*(2nd ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Szydlik, J. E., Kuennen, E. W., Belnap, J. K., Parrott, A. L., & Seaman, C. E. (2011). Changing Mathematical Sophistication in Introductory College Mathematics Courses. In *Proceedings of the Fourteenth SIGMAA on Research in Undergraduate Mathematics Education Conference*. Portland: OR.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*, 101-119.

DIFFICULTIES IN USING VARIABLES – A TERTIARY TRANSITION STUDY

Ileana Borja-Tecuatl, María Trigueros, Asuman Oktaç
Dept. of Mathematics Education, CINVESTAV-IPN, Academy of Mathematics, UACM;
Academic Dept. of Mathematics, ITAM; Dept. of Mathematics Education CINVESTAV-IPN

In this paper we describe the results obtained from a diagnostic instrument that was applied to 25 students of the Universidad Autónoma de la Ciudad de México (UACM) at the beginning of an Algebra and Analytic Geometry course. The diagnostic instrument was the first part of the study we currently perform to analyze the possible impact that failing to understand the uses of variables may have in understanding systems of linear equations. A brief description of the main research project is presented to frame the results of the diagnostic instrument into context.

Key words: 3 Uses of Variable Model, Linear equations, Systems of linear equations.

This study has been held in a public university located in Mexico City that does not apply an admission exam; the only requisite is a High School Diploma. If the number of applicants is larger than the number of places available, the decision is taken by means of a lottery in front of a public notary. Furthermore, the students that this university receives in most of the cases have interrupted their studies from high school to university, come from the least favoured zones in Mexico City and cannot commit themselves to be full-time-students, since they have the need to work either part-time (4 to 6 hrs. a day) or full-time (8 to 10 hrs. a day) to support their families.

Students in this study are registered in an engineering program, where they require a strong background in mathematics. Taking into account results obtained in many research studies about the importance of a good understanding of variable to be able to succeed advanced mathematics courses (Dorier, Robert, Robinet & Rogalski, 2000; Gray, Loud & Sokolowski, 2005; Trigueros & Jacobs, 2008; Trigueros, Oktaç & Manzanero, 2007; Trigueros & Ursini, 2008; Ursini & Trigueros, 1997, 1999, 2001), it was decided to study the understanding of variable of these students in order to design and teach a remedial course focusing in algebra and in particular in the flexible use of variables to help them cope with the mathematics courses they need to take.

Our teaching experiences in this university, prior to the beginning of this study, suggested that those students did not have a clear understanding of the different uses of variables, nor how to deal with them, and were not able to take advantage of the official courses in engineering, as demonstrated by the failure to pass their official[1] courses. This observation agrees with the one that Trigueros and Jacobs (2008, p.105) recall, where "according to Trigueros and Ursini (1999, 2001, 2003) a well-developed understanding of algebra

[1] We refer to official courses as those that constitute the engineering programs offered in the University. The University also offers remedial courses, prior to the official ones, as a help to overcome the difficulties shown by students in the diagnostic examination that the Institution applies to all new students. Up to August 2011, the diagnostic examination was used by the Institution to decide which students could enrol in the official courses and which needed remedial help before enrolling in them. Every year the majority of the students were assigned to the remedial courses, so from August 2011 all students are registered to the remedial courses independently of their results. Though it was always expected that students would commit to the remedial instruction in the practice this rarely happens given that the developmental course is not mandatory or certified by the Institution.

necessitates the ability to differentiate among the three uses of variable and to flexibly integrate their uses during the solution of any problem".

Within this context, the main objective of our study was therefore the identification of students' specific difficulties when using variables, to be able to design and apply a didactic treatment based on the 3 Uses of Variables Model (3UV Model; Ursini & Trigueros, 1997, 2001) that would, on one hand, allow us to make the best out of class-time to help students improve their understanding of variable (since students in general have very little or no time to do homework) and, on the other hand, assess at the end of the didactic treatment how students' conception of variable is related to achieving a correct mental construction of the solution to simultaneous linear equations, from the linear algebra perspective.

In this university, linear algebra is the first course of the engineering programs that belongs properly to the tertiary level of education. Previous research has suggested that the lack of understanding of the different uses of variables can contribute to students' difficulties in understanding Linear Algebra (Dorier, Robert, Robinet & Rogalski, 2000; Trigueros, Okta□ & Manzanero, 2007), particularly to determining the solution to a system of linear equations (Trigueros, Okta□ & Manzanero, 2007). Our research project intends to see if explicitly teaching variables to students contributes to their developing of a better understanding of Linear Algebra concepts.

The content of this paper concentrates in the design and results of the instrument we used to study students' understanding of the uses of variable. In the next sections we describe the theoretical framework used to design the instrument and analyze the results obtained.

Methods and Theoretical Framework

The purpose of the diagnostic instrument was to identify how the students deal with the different uses of variables. Naturally the 3 Uses of Variable Model (3UV Model) served as a theoretical reference to design the instrument and analyze students' possible and actual responses.

The 3UV Model is a theoretical framework proposed by Ursini and Trigueros (1997, 2001) as "a basis to analyze students' responses to algebraic problems, to compare students' performance at different school levels in terms of their difficulties with this concept, and to develop activities to teach the concept of variable" (Trigueros & Ursini, 2008, p.4-337). The 3UV model takes into consideration the three most frequently uses of variable present in elementary algebra: specific unknown, general number and variables in functional relationship. Its authors emphasized aspects corresponding to different levels of abstraction at which each one of the uses of variable can be handled. These aspects are described in the following paragraphs:

According to the 3UV Model, the understanding of variable as unknown requires to: recognize and identify in a problem situation the presence of something unknown that can be determined by considering the restrictions of the problem (**U1**); interpret the symbols that appear in an equation, as representing specific values (**U2**); substitute to the variable the value or values that make the equation a true statement (**U3**); determine the unknown quantity that appears in equations or problems by performing the required algebraic and/or arithmetic operations (**U4**); symbolize the unknown quantities identified in a specific situation and use them to pose equations (**U5**).

The understanding of variable as a general number, according to the 3UV Model, implies to be able to: recognize patterns, perceive rules and methods in sequences and in families of problems (**G1**); interpret a symbol as representing a general, indeterminate entity that can assume any value (**G2**); deduce general rules and general methods in sequences and families of problems (**G3**); manipulate (simplify, develop) the symbolic variable (**G4**); symbolize general statements, rules or methods (**G5**).

As the 3UV Model considers it, the understanding of variables in functional relationships (related variables) implies to be able to: recognize the correspondence between related variables independently of the representation used (**F1**); determine the values of the dependent variable given the value of the independent one (**F2**); determine the values of the independent variable given the value of the dependent one (**F3**); recognize the joint variation of the variables involved in a relation independently of the representation used (**F4**); determine the interval of variation of one variable given the interval of variation of the other one (**F5**); symbolize a functional relationship based on the analysis of the data of a problem (**F6**) (Trigueros & Ursini, 2008).

To design the instrument's items, a literature review was performed to identify the obstacles that students face in understanding systems of linear equations in school algebra. Findings indicate that students have problems to understand concepts that are prerequisite to the understanding of systems of linear equations, which are actually prerequisites to it (Malisani & Spagnolo, 2009; Panizza, Sadovsky & Sessa, 1999). These results led us to review what ate the common difficulties when working with variables described in the literature (Fujii, 2003; Phillip, 1992; Sfard & Linchevski, 1994; Trigueros & Jacobs, 2008; Wagner, 1983), and those related specifically to linear equations in one variable (Herscovics & Linchevski, 1994, 1996) and linear equations in two variables (Malisani & Spagnolo, 2009; Panizza, Sadovski & Sessa, 1999). From this review a set of 21 abilities were extracted that we considered important to verify regarding the preliminary concepts for the study of simultaneous linear equations and that previous studies suggest as common difficulties among students in secondary level of education.

The items of the diagnostic instrument were then designed to provide information about students' performance in three aspects: 1) the 21 abilities identified as a result of the literature review, 2) the different uses of variables according to the 3UV Model and 3) identification of a change of use in the variable in some problem situations.

The instrument consisted of 28 items. Some items were purposely designed to include more than one use of variable along the solution of the problem presented in the item to analyze if the students were able to recognize and work with the different uses of the variables needed in their solution.

In order to analyze students' solutions to the instrument we prepared two tables, one in terms of the 3UV Model and the other one in terms of the abilities selected from the literature review. Each of these tables included a column for each item of the instrument and a row for each of the abilities to be considered. A 1 was assigned, if there was clear evidence in the student's response for that ability; 0 indicated that there was not clear evidence for that ability; and the cell was left empty if there was no response. Both the tables were filled for each of the participating students.

The participating students were those who attended the first two 75-minute-sessions of the course on Algebra and Analytic Geometry, one of the official courses in the first semester of the engineering programs, taught by one of the authors the first semester of 2011 at the UACM, where the instrument was applied. About half of the 25 students that participated had enrolled in the university the previous semester and many of them had not attended the remedial instruction; the rest of the students had been registered in the University for at least a year but the students were either taking part in that course for the first time or repeating the course. These conditions were adequate to verify if there was a significant difference in the understanding of algebra between students who had taken or not the remedial instruction and those who had taken some courses in the University for at least one year. The results of the analysis are described in the following section.

Results of the diagnostic instrument

In this section we describe the results for each of the uses of the variables. For each of them, we present a graph that shows how many students out of 25 gave evidence of each of the respective abilities included in the 3UV Model. The information is presented in bar-graphs; each bar corresponds to one specific item in which the ability was tested.

For those cases in which the information coming from the problematic abilities taken from the literature review gives a complementary insight of the difficulties that students show in working with the variables, we also provide that information. In many cases the information was a confirmation of what we found with the 3UV Model.

The use of variables as general number

Six items were considered to test students' use of variables as general numbers: items 2, 7, 10 11, 15 and 26. They are presented in the following paragraphs, followed by their respective analysis.

Exercise 2: Give three examples in which you assign values to z in the expression $2 + z + z + z + 3z - 5 - 2z$ and simplify your result.

The purpose of this exercise was to find out whether students would first manipulate the expression (G4) and later assign the values as an example of the different values that z could take in that expression (G2) and finally give the result, or students would first assign the values and then simplify. For the purpose of analysis, if students assign different values to each z appearing in the expression before simplifying, they would show a lack of consistency of meaning throughout a single context (Wagner, 1983; Fujii, 2003). If students first assign a value to z and give different examples for that value, it would be considered that they have developed G2, but not G4. If they would simplify before providing three different examples of z, it would be considered that they are able to do both G2 and G4. If students simplify and then transform the expression into an equation to find the value that would satisfy it or state that they cannot know the values for z because they don't have information about z, it would be considered evidence for G4 but not for G2.

Exercise 7: Write a mathematical expression that describes that 5 is added to a number.

The purpose of this exercise was to verify whether students would consider a number as a general entity that can take any value (G2) or if they would consider it a specific representation of a fixed quantity that they could assign, as well as verify how they would symbolize the addition (G5). If students consider a general number and add it to 5 it would be considered evidence for both G2 and G5; if students write the addition of 5 to a specific number, it would be considered evidence for G5 but not for G2, since in the question students needed to consider the expression 'a number' as a general number that can take any value, and assign one of them that would be added up to 5 to later write that addition. In such case, since they only implicitly thought of a general number, they do not show clear evidence for G2. Any expression that involves an addition, but does not involve 5 as one of the addends was not considered evidence either for G2 nor for G5. If a symbol for a general number was given in relation to 5, but not in a sum, that was considered evidence for G2 but not for G5.

Exercise 10: Simplify the following expression $3 + \boldsymbol{x} - 4\,\boldsymbol{x} + \boldsymbol{x} - 2(\boldsymbol{y} - \boldsymbol{x} + 2) + 5 =$.

The purpose of this exercise was to verify whether students would recognize the use of the variables as general indeterminate entities that can assume any value (G2), or if they would change the use the variables transforming the expression into an equation. We also wanted to know if students were able to manipulate the variables correctly (G4). If students manipulate the variables correctly and keep the expression as an open expression, it was considered evidence for G2 and G4. Manipulating them correctly but transforming the expression into an equation was considered evidence for G4, but not for G2. Not manipulating the variables correctly would account as evidence that the student did not show G4.

Exercise 11: Answer the following questions: In the expression $4 - \boldsymbol{x} - \boldsymbol{y}$, a) Can $\boldsymbol{x}$ equal 3.1416? b) Can $\boldsymbol{x}$ and $\boldsymbol{y}$ hold the same value? Give an example and substitute it. c) Can $\boldsymbol{x}$ or $\boldsymbol{y}$ be negative? Can they be positive? d) How many values do $\boldsymbol{x}$ and $\boldsymbol{y}$ represent? Give examples.

The purpose was to explore whether students consider it possible that $\boldsymbol{x}$ and $\boldsymbol{y}$ represent any value at once (G2) and if among the possible values they considered, negative, non-integer values or the same value for $\boldsymbol{x}$ and $\boldsymbol{y}$ simultaneously were considered. If one of the questions was answered negatively or stating that they represent only one value, it was considered as not giving evidence for G2.

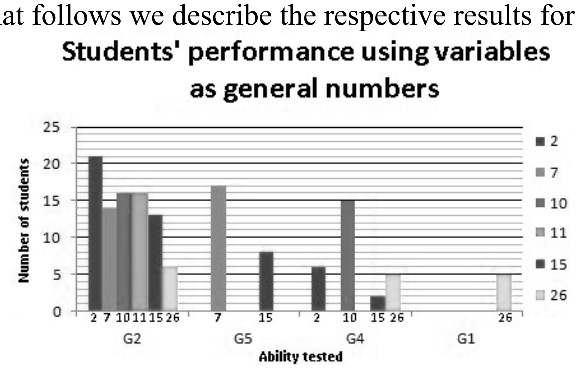
Exercise 15: Simplify this expression $\frac{3}{a^2} - \frac{5}{a^2 b} + \frac{4}{ac} =$.

The purpose of the question was to find out if students could work with variables as general entities without having to assign values (G2) manipulate them as being denominators in a sum of fractions (G4) and symbolize the result of their operations correctly (G5).

Exercise 26: Write down all the operations you need to transform $16\boldsymbol{h}^2 - 24\boldsymbol{h} + 9$ into its corresponding squared binomial. Remember, a perfectly squared trinomial $\boldsymbol{a}^2 - 2\boldsymbol{ab} + \boldsymbol{b}^2$ can be written as the squared binomial $(\boldsymbol{a}-\boldsymbol{b})^2$.

The purpose of the exercise was to know whether students recognized the trinomial as similar to the one exemplified as a general representative of a special family of trinomials that had been introduced since secondary school (G1), and if so, if they could manipulate the open expression without transforming it into an equation (G2) to put it in the corresponding form as a squared binomial (G4). If students substitute values for $\boldsymbol{a}$ and $\boldsymbol{b}$, either tacitly or explicitly, that would allow them to transform the expression, it would be considered evidence for G1, but it would only be considered evidence for G4 if they explicitly stated the operations to transform the trinomial into the squared binomial. If students didn't need to transform the expression into an equation, that would be considered evidence for G2. When the substitution was achieved, but the student needed to transform the expression into an equation to be able to operate with it, it was not considered as evidence for G2, since what was shown was an interference of the unknown with other uses of variables (Gray, Loud & Sokolowski, 2005; Malisani & Spagnolo, 2009; Philipp, 1992).

The results show that student's performance while dealing with exercises that involve this use of the variable is only relatively favorable for G2. In average, 14 students gave clear evidence that they could interpret a symbol as representing a general, indeterminate entity that can assume any value in all the items for which the ability was tested. For the rest of the abilities tested (G1, G4 and G5) the performance was not so good. For all the abilities, the students' performance strongly depended on the complexity of the exercises: the higher the complexity of the exercise, the lower the performance. This information is shown in the following graph. In what follows we describe the respective results for each item.



The results show that the students assign the same value to all the appearances of the variable in the expression of exercise 2 and are able to give different examples for the value of z. However only 6 students out of the 25 that answered the item would first manipulate the expression and then provide the values that z can take. It was common that students that would first simplify would then conclude that 'z can take any value because it is a variable' (4 out of 6 students). Such an answer was considered evidence for both G2 and G4, independently of whether the students provided examples of the value of z or not.

Regarding exercise 7, only 17 students out of the 25 that answered the item were able to express an addition of 5 to a number, either general or specific, but only 14 students did consider a general number in the expression they wrote. The main error in this exercise was that students wrote additions which gave 5 as the sum.

Only 15 students out of 23 that answered item 10 were able to manipulate correctly the expression and 16 out of the 23 didn't transform the expression into an equation. This means that about 1/3 of the students that answered this question have a tendency to react to an equal sign as describing an equation and, in general, they would complete it by writing 0 to the right side of the equation. We also noted that students had difficulties to manipulate the expression correctly, either because they couldn't group like terms, a difficulty described by Herscovics and Linchevski (1994, 1996) or because they did group like terms correctly but had difficulties performing the required arithmetic operations, another difficulty reported by Herscovics and Linchevski (1994).

23 students answered item 11. From them, 16 considered that x and y represent any value, either positive, negative or non-integer, and that they could be the same, but though students answered yes to the questions, the examples they provided show that, in general, they are used to give different values to different variables and that the values they assign are natural numbers (not even integers). This means that although they can agree to the fact that a symbolic variable can represent many values, either positive, negative or non-integer, in practice they constrain their answers to the natural numbers and use different values for different variables. This was also verified in the examples they provided for exercise 2, where only two students gave negative values among their examples and only four included 0 as a possible value, none student gave non-integer values as example of the values that a general number can take.

In exercise 11, one student wrote "yes, y can be negative but only if you write a positive value for y", meaning that he was actually not thinking of **y** (nor **x**) being negative, but –**y**, as it is written in the exercise. This makes us believe that some other students might have answered yes to the fact that the variables could be negative, but were indeed not meaning it as a value for **y** but for –**y**, what would have accounted for less students giving evidence for G2, but we couldn't know from their answers because they were too general as "yes, since they are variables, they can take any value".

Among the students that didn't give evidence for G2, some of them answered yes to the questions and also stated that the values could be negative, equal or non-integer but argued that "yes they can, but it depends on the values that balance the equation", showing that in practice they refer to the variables as being related variables or unknowns, and not general numbers.

The results show that exercise 15 proved to be very difficult for most of the students. Only 14 of 25 tried to solve it; 13 of them didn't need to assign specific values to operate the expression, but only 3 of them manipulated the variables in a way that corresponded to a sum of fractions. Not all of them could find a correct common denominator. 8 were able to symbolize correctly the results of the operations they were performing while the other 5 students couldn't even symbolize what they were meaning. In such cases the most common mistake was to consider that when a variable is multiplied by a scalar, the exponent also becomes a coefficient, or that terms can be added up indistinctively of whether they are like terms or not.

In this exercise it was interesting to notice that even though only three students were able to find a correct common denominator, they actually did it by identifying the structure of the expression. That was not the case when, in a previous arithmetic question in the instrument they were asked to solve $\frac{3}{2^2} - \frac{5}{7(2^2)} + \frac{4}{3(2)} =$. This sum has exactly the same structure as that in exercise 15, where the denominators are letters. In the arithmetic question none of the students was able to recognize the structure so they didn't use it to find the common denominator. All of them expanded the squares, multiplied the numbers and then found the common denominator for the resulting quantities. We suspect that the fact that the expression did not include numbers that students can operate with, they are not compelled to expand the terms and therefore, if they know what to do, they directly apply a rule to find a common denominator, but in the presence of numbers they are not used to consider the structure of the expression. This hypothesis needs to be verified, but it was surprising to notice that none of the students actually gave importance to nor used the structure present in the arithmetic question.

Only 6 students answered to exercise 26, and it was not due to lack of time, since students actually gave back their responses before the end of the second session. This shows that the question was very challenging to them, so much that they preferred to skip it. The results show that none of the six students that answered the item transformed the expression into an equation and that only one of them couldn't transform it into the corresponding binomial. Also, 5 of the 6 students were repeating the course and had taken Algebra and Analytic Geometry the previous semester. The other student was taking the course for the first time, had not taken the remedial instruction and when trying to operate the expression he gave evidence of misunderstandings about operations with variables; he wrote "$16\boldsymbol{h}^2 - 24\boldsymbol{h} = 8\boldsymbol{h}$" and "$8\boldsymbol{h} + 9 = 17\boldsymbol{h}$". This mistake was also found in the answers to question 15.

As it has been shown, students do not know how to operate variables; they gave evidence of failing to use the basic rules that have to be followed to accurately operate with them. They gave some evidence of difficulties in differentiation the notion of unknown and the use of variables as general numbers, and that they performed confidently and correctly only for very simple familiar tasks. In conclusion, their use of variables as general numbers has to be strongly nurtured and fostered.

The use of variables as unknowns

Four items were used to investigate the use of variables as unknowns. They are presented in the following paragraphs followed by their analysis:

Exercise 3: How many and which values satisfy the expression $3 + \boldsymbol{a} + 6 = 25$? Write down how you found the answer.

Exercise 4: How many and which values satisfy the expression $3 + \boldsymbol{a}^2 + 6 = 25$? Write down how you found the answer.

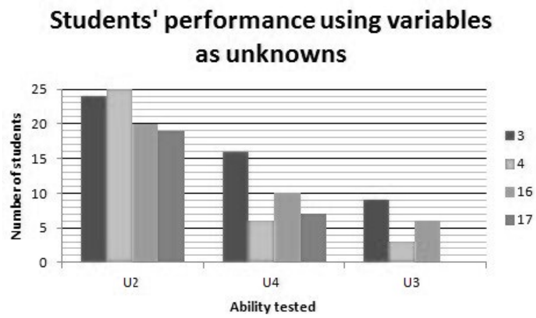
Exercise 16: Find all the solutions to the equation $\boldsymbol{z} + \boldsymbol{z} + \boldsymbol{z} + 7\boldsymbol{z} - 2\boldsymbol{z} = 72$.

Exercise 17: Solve the equation $-3 + 3\boldsymbol{w} + 6 - 6\boldsymbol{w} + 23 = 10 + 7(1 - \boldsymbol{w}) - 7 - 4\boldsymbol{w}$.

For all the exercises we applied the following criteria to analyze the responses: if students give answers that revealed that they were looking for specific values, it was considered evidence for U2, even if the values they got were not right. If they use arithmetic procedures, which require verifying the solution, or if they verified it although they used algebraic procedures, we would consider it to be evidence for U3. If students determine the unknown quantity by algebraic procedures, it was taken as evidence for G4. Not succeeding in determining the unknown quantity, although students used algebra, was not considered evidence for G4. In reality the 3UV Model considers U4 as the ability to determine the unknown quantity either by performing algebraic or arithmetic operations, but our main purpose while designing the items was to find out how many students actually use algebra while solving equations, even if they are presented to simple equations. To differentiate between students that used algebra and those who didn't, we adapted the criteria to what we have explained. For Exercise 4, evidence for U3 and U4 was considered only if students stated both the values that satisfy the equation.

When presented with exercises that involve the use of variables as unknowns, students seem to be most at ease with interpreting the symbols that appear in equations, as representing specific values (U2), compared to determining the unknown quantity that appears in equations (U4) or substituting to the variable the value or values that make the equation a true statement (U3).

These results are condensed in the following graph that we will explain in the following paragraphs in relation to the specific items involved.



Knowing that 25 students answered the first two items and 20 answered the last two, the graph tells us that all students with except of one for exercise 3 and another one for exercise 17 considered the variable to have specific values. In exercise 3, one student tried to solve by algebraic means, and actually got the correct value, but when answering which values satisfied the equation she wrote "any value". In Exercise 17 other student made as well use of algebra, but lost track of the '=' sign, transforming the equation into an open expression. It was noticeable that the student actually didn't know where he was going by manipulating the expressions and only for the first two items he was certain that the variables given in the expressions were unknowns.

Surprisingly, most of the students applied algebra (20 students for item 3; 16 for item 4; 14 for item 16 and 20 for item 17) even while solving the first two equations, but only in the case of equation 17 all the students tried to solve it with algebra. The graph shows that only 16, 6, 10 and 7 students succeeded in solving the respective equations by algebraic means. Students' main obstacle to find the solution was arithmetic mistakes, among them not adding correctly the coefficients of like-terms, ignoring minus signs, not balancing the equations correctly while trying to group like-terms on different sides of the equation and ignoring the priority of operations. All those difficulties were reported by Herscovics and Linchevski (1996). The results show as well that there are some students that tend not to use algebraic procedures if it is not that difficult to solve the equation by arithmetic operations and that only in the case of equations presenting the variable on both sides of the equation they used algebraic operations from the beginning in the solution of the problem. This situation was reported as well by Herscovics and Linchevski (1996).

From all the students that attempted to solve by algebraic means, only four students verified the solution for item 3, and only two students did it for items 4 and 16. No one verified the solution for item 17. Sfard and Linchevski (1994) have reported on this problem, where students misunderstand the solution to an equation with the procedure to solve it, as if solving it would mean to find a value through the algebraic procedure and not to find a value that satisfies it.

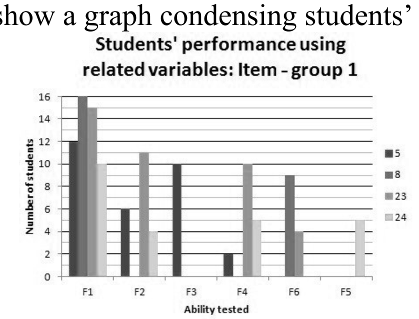
Knowing how many students tried to verify their solutions the numbers shown in the graph for U3 must be read as the following: the five students that didn't use algebra for item 3 succeeded in finding a solution, only one of the nine students that didn't use algebra for item 4 succeeded in giving the two values that satisfy the equation and only four of the six students that didn't use algebra for item 16 succeeded in finding the correct value that would make the equation valid.

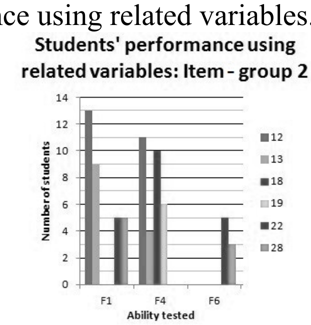
The results show that students find it difficult to consider negative values as solutions to an equation. In item 4, only 7 of 25 students considered the value –4 as solution for the equation. In item 17, some students that completed the procedure to solve the equation ignored the minus sign and gave the solution as if it was positive. Some other students left the solution procedure up to a point in which they avoided concluding, as if they had realized that the solution would give negative and preferred not to state that. Since students didn't verify their answers, they didn't realize the mistake of considering the solution to be positive.

In items 3 and 4, where students were asked to specify how many solutions the equations have, students omitted to indicate explicitly the number of solutions and left the teacher to interpret from what they wrote. It has been reported by Trigueros and Ursini (2008) that students tend to forget what the purpose of the problem is, and that they rarely turn their attention back to the questions to check whether finding a solution was enough for solving the problem posed or if they would still need to do something else; our students share this problematic situation.

The use of variables in functional relationship

We included 10 items involving related variables that we organized in two different groups. The first group has four items (5, 8, 23 and 24) and involves all of the abilities included in the 3UV Model for related variables. The second group has six items that focus on F1, F4 and F6. These items are related in pairs to one another (12, 13; 18, 22; 19, 28). We present two graphs, one for each of the groups, and describe the results separately. At the end we give a general conclusion about students' performance with the use of related variables and show a graph condensing students' global performance using related variables.





Exercise 5 was "Give values to ***a*** and ***b*** as to make the expression 3 + ***a*** + 6 = ***b*** a valid statement. Write down your operations. How many values could be given for each variable?" The purpose of the item was to identify to which level of abstraction, according to the 3UV Model, students understand the given expression. First, we wanted to find out if students recognize the equation as one expressing related variables (F1) or if their understanding of the expression is that of "process equals result", a conceptual limitation described by Sfard and Linchevski (1994). Evidence for F1 was considered if the student gives at least two different examples in which he/she assigns values to either ***a*** or ***b***, or if the student manipulates the expression to transform it in one where ***a*** is considered the dependent variable. The manipulation of the expression to invert the dependency of the variables was also considered evidence for F3. Another evidence for F3 would be that students start assigning values to **b** and found the corresponding values to **a**. Evidence for F2 would be that students assign at least two different values to ***a*** to determine the corresponding values of ***b***. Evidence of F4 was considered if students give insight that they understand that the variables are linked by the equation and that changes in one affect the other one, independently of which one is changed. If students give one single example stating the value of ***a*** and ***b*** (in that order), it was considered as an understanding of process equals result, because to consider ***a*** as an independent variable, they had to show a variation in ***a***.

The results show that only 12 students could give evidence for F1, that ten changed the dependency of the variables (F3), that six recognized the dependency of variables as it was stated (F2) and that only two expressed a correct understanding of joint variation (F4). 21 students answered to the item, but only 12 students were able to recognize the correspondence between the variables and most of them needed to manipulate the expression to become aware of the dependence. The results reveal that of 25 students that participated and 21 that answered the question, only 2 have a complete understanding of the most important levels to which related variables need to be understood to benefit from an Algebra and Analytic Geometry course.

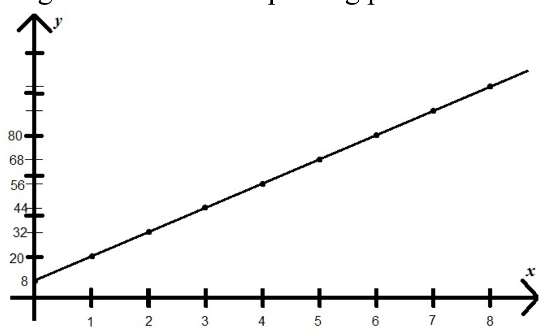
Exercise 8 was "If the cost of each kilogram of coffee is 88 pesos, write down an expression that would allow calculating the cost, ***c***, of ***n*** kilograms of coffee." The purpose of the exercise was to detect whether students are able to identify the relationship between the variables (F1) and to write it down in a symbolic way relating the two variables involved (F6). We considered it evidence for F1that students try to calculate the cost either for specific values or for general numbers. If students succeed in symbolizing the corresponding equation of the correspondence between ***n*** and ***c***, we would consider that as evidence for F6.

The results shows that only 16 students out of 23 that answered the item were able to identify the correspondence between the involved variables, but only nine were able to symbolize this proportional relationship as an equation in related variables.

Exercise 23 was: "In the local market there is a lady, Juanita, who sells fresh orange juice. The liter of juice costs 12 pesos, but she charges a fix amount of 3 pesos for her service, independently of how many liters of juice a client buys. Use that information to a) calculate how much a client would pay including the service if he/she buys 0.5, 1 or 2 liters of juice, respectively; b) organize the information into a table and graph the corresponding cost for the different amounts of juice; c) write down an expression that would allow Juanita to know how much she should charge in total (including the service) to someone that wants ***k*** liters of juice." The purpose of the exercise was to find out if the students recognize that there is a correspondence between the involved variables (F1) by attempting to calculate the cost for each of the amounts of juice given in the problem and try to describe the relationship between the variables even if not correctly. If students calculate correctly, that would be considered evidence for F2. If students represent the information in to a table and graph that would describe how the variables varied according to the way they calculated, that would be considered evidence for F4. If students symbolize a general expression to describe the relationship between the variables according to the information given, that would be considered evidence for F6.

The results show that all 15 students that answered the questions could give evidence for F1, but only 11 could find the correct values for the dependent variables, 10 could represent the information in a graph and table and 4 could symbolize the relationship between the variables. Notice that while similar amounts of students could recognize the correspondence between the variables in exercises 8 and 23 (16 vs. 15, respectively), the number of students that could symbolize a linear relation decreased considerably (9 vs. 4, respectively). The students that didn't succeed in calculating the values for the dependent variable or symbolizing the correspondence considered the correspondence to be proportional, a problem that was reported by Bardini and Stacey (2006).

Exercise 24 was "The following graph represents the total amount to be paid, ***y***, at a parking lot in a mall located to the south of the city, accordingly to the amount of time, ***x***, that a car uses the parking place. Analyze the graph and answer the following questions: a) what amount has to be paid for using the parking place 2, 2.5 and 3 hours, respectively? b) How would you interpret the value of $\boldsymbol{y}$ when $\boldsymbol{x} = 0$? c) Angelica's car was parked for 4 hours; she estimates to pay between 40 and 45 pesos. Is her estimation right? How much would she need to pay for the time she used the parking place? d) If $\boldsymbol{y}$ can vary between 20 and 68, for how long could a car use a parking place?"



The purpose of this exercise was to verify if students can extract basic information from a graph about the way in which the corresponding variables are related and to find out how students interpret the value of $\boldsymbol{y}$ for $\boldsymbol{x} = 0$. We considered it evidence for F1, if at least two of the requested values of $\boldsymbol{y}$ were correct in question a). If the answer for $\boldsymbol{y}$ to question c) was correct, we considered it evidence for F2. We considered it evidence for F4, if students would relate their answer to question c) to the interval given there or if they gave an interval for question d), even if not correct. Only if the interval for question d) was correct, it was considered evidence for F5.

The results show that ten out of the 11 students that answered the item could give evidence for F1, but only four were able to extract correctly from the graph a specific piece of information to give evidence for F2. None student related the interval given in question c) to justify why Angelica was not right in her estimation, so the information coincides for F5 and F4, meaning that the same five students that gave an interval for question d did it correctly. Only 2 students were able to interpret the graph completely as to give evidence for all the abilities involved and find all the information that was required, including a correct interpretation of the value of $\boldsymbol{y}$ when $\boldsymbol{x} = 0$. One student interpreted $\boldsymbol{y}$ as the cost to enter the parking lot, and the other one as a fix maintenance (service) amount. Besides them only two other students tried to interpret and only one did it correctly, interpreting as the amount to pay as one enters the parking lot.

In conclusion, the results for the first group of items show that not even for very common familiar situations students are able to determine values of related variables if one or the other is given. They also cannot extract completely the information presented in a graph nor represent in a graph the variables they are relating. All these abilities are required to benefit from the contents of the course in Algebra and Analytic Geometry and other courses in the engineering programs.

From the items included in second group we present exercises 18 and 22 and 19 and 24; we will explain them in pairs since they are related to each other.

Exercise 18 was "What geometrical object does equation $\boldsymbol{x} - \boldsymbol{y} + 9 = 5$ represent?" The purpose of the exercise was to know whether students recognize the equation as an algebraic representation of a straight line that crosses the $\boldsymbol{y}$ axis at 4 and has slope 1. We also wanted to find out how students determine the geometric representation. Evidence for F4 would be considered if students relate correctly the equation to a straight line of specific characteristics, since the variables vary jointly along that particular line. In the process of analyzing the information, we had to reduce our consideration for evidence to answers that recognize the equation as a straight line even if its characteristics were not stated. This kind of response do not offer clear evidence that the student understand the joint variation of the variables, but none of the students attempted to describe what particular characteristics the line has and only two students, out of 16 that answered the item, gave an argument for their answers; only one student related the equation to "$\boldsymbol{y} = \boldsymbol{m}\text{x} + \boldsymbol{b}$", but he didn't use this structure to specify the characteristics of the line. The instructions to the instrument requested to always justify the answers; for this item and item 19 students didn't follow the instructions.

Exercise 22 was "Give an example of an equation representing a straight line." The purpose of this item was to control random correct answers for exercise 18. We considered it evidence for F1 if students gave an equation in related variables, even if not correct for a straight line and we considered it evidence for F6 if students gave a correct equation for a straight line.

The results show that ten students claimed to recognize the equation in exercise 18 as representing a straight line and that five students succeeded in giving an example of an equation to a straight line. By crossing the information from both the exercises, we found out that of the 10 students that claimed to recognize the equation in exercise 18 as representing a line, only three students also gave a correct example of an equation. From the rest of these 10 students, six didn't answer and one gave an open expression. The other two correct examples were given by students that didn't identify the equation of exercise 18 as representing a straight line. One of those students stated that the equation represented an ellipse and the other one that the equation represented a circumference. Among the mistaken representations attributed to the equation we found four proposals for a triangle, one for an ellipse, and one for a circle; one of the triangles and one the circles were proposed by students repeating the course. None of the 16 students that answered item 18 tried to generate a graph to support or verify their answers.

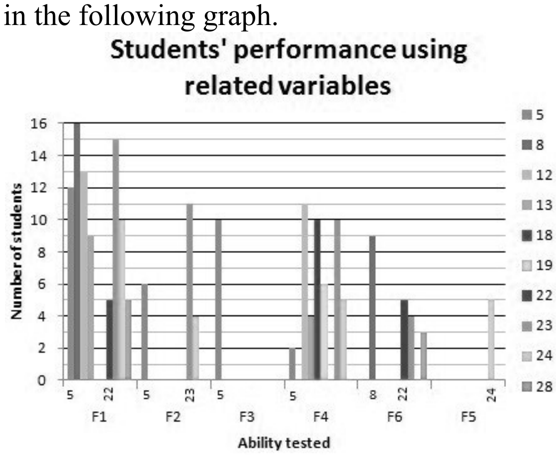
Exercise 19 was "What geometrical object does equation $\boldsymbol{x}^2 + \boldsymbol{y}^2 = 9$ represent?" The purpose of the exercise was to know whether students recognize the equation given as an algebraic representation of a circumference centered in the origin with radius 3. We also wanted to find out what means students use to determine the geometric representation. Evidence for F4 would be considered if students were able to correctly relate the equation to a circumference of specific characteristics, since the way variables vary jointly is along that particular circumference. Though, we had to be content in the process of analyzing the information to consider as evidence just the fact that the students related the equation to a circle or circumference, without specifying its characteristics. We recognize that is very arguable, but we accepted to process like that given the fact that only one student, out of 14 that answered the item, gave an argument to the answer and actually specified the characteristics; another reason for it is that though in strict sense a circle and a circumference are not the same, students seem to not be aware of that difference. Among the students referring to the geometrical representation as a circle, there were four students that had taken the course before and two of them were the students showing evidence for all four abilities tested in exercise 5.

Exercise 28 was "Give an example of an equation representing a circumference with radius 3." As a way to control random correct answers for exercise 19, we asked to give an example of an equation of a circumference of radius 3. Since it is open to any kind of equation representing a circumference of radius 3, we would consider it evidence for F1 if students would actually give an equation in related variables, even if not correct for the requested data. We would consider it evidence for F6 if the given equation was a correct equation for a circumference of radius 3.

The results show that six students claimed to recognize the equation of exercise 19 as to represent a circumference (or circle) but only three students were able to exemplify a correct equation, though five students proposed an equation in related variables. Of the six students that claimed to recognize the equation to be a circumference or circle, only two could give a correct example of a circumference with radius 3. The other four students didn't answer exercise 28. The third correct example came from a student that had not answered item 19. The reasons for the failure of the other two proposals were in one case not giving radius 3 but its root, and in the other not giving a quadratic equation, but linear. Among the mistaken representations attributed to the equation in exercise 19 there was one rectangle, two parabolas, two areas of a square, one hyperbola, one "curve", and one straight line. Students that had taken the course before suggested the parabolas and hyperbolas. None student tried to generate a graph to support or verify their answers.

Besides the detected difficulties that have been explained, these four exercises also show that students do not consider changing between the symbolic, tabular and graphical representation to verify or support their answers. The study of Analytic Geometry and Calculus, among other courses in engineering requires the flexible transition between different representations to understand their contents. Students present several obstacles to benefit from the courses they have to take to become engineers.

In conclusion, the results reveal that with exception of some cases in which the context is familiar to the students, they do not know how to act when confronted with equations in related variables. In general for all the exercises that involved this use of the variables, they only perform relatively well for recognizing the correspondence between related variables independently of the representation used (F1). In average for all the tasks involving related variables, 11 students recognized the correspondence between the related variables; while only 7 students could determine the values of the dependent variable given the value of the independent value (F2) and 10 students could determine conversely (F3), only 7 could recognize the joint variation of the variables involved in a relation independently of the representation used (F4), only 5 students could determine the interval of variation of one variable given the interval of variation of the other one (F5) and also only 5 students could symbolize a functional relationship based on the analysis of the data of a problem (F6). This information is condensed in the following graph.



Results for transitions in the uses of variables

Students have really strong difficulties to adapt to a change in the use of a variable along the solution of a single problem. In general they would start by manipulating the expressions, without a clear strategy of what they plan to do, and after wandering for a while they would either give up the procedures without summing up or stick to one trial and follow it until they somehow would end up with a response that they never verified to be a solution.

A representative exercise for this section was "How many solutions does the equation $2\boldsymbol{y} - 3\boldsymbol{x} = 0$ have? Exemplify your answer." This exercise was answered by 18 students, of them 7 used the equation to express a explicit dependent relation, but only one actually substituted a fix value to the independent variable to solve the equation and generate examples of the solutions to the equation. In general what these students did was to express both the variables as dependent variables, respectively, and then either be blocked because they couldn't adapt to assign values or they tried to solve a system of equations, that in some cases –due to mistakes- actually led them to what they claim to be a unique solution. The rest of the students mistook the equation as to have one unique solution (0, 0).

There were four exercises involving a change of variable in the instrument. Two were of the type we just showed, one was related to exercise 24, that means there were three exercises involving a change from related variables to unknowns, and the last exercise involved a change from general variable to unknown by substituting a perfectly squared trinomial by its correspondent squared binomial to solve an equation. The inability to adapt to the required changes in use of the variables was present in all four exercises.

Concluding Remarks

As we have shown in the previous analysis, students that participated in this study are far from having a clear understanding of how variables are operated and show several weaknesses regarding the abilities that would help them identify the different uses of variables and adapt to changes in the use of the variables along one same problem. The reasons for student's weaknesses in identifying and using correctly the different uses that variables can take are essentially due to common difficulties reported in the literature. No significant difference was found between the types of mistakes that students starting university and students that had been in university for at least a year made, with except of those specifically expressed during the past sections. This reveals that not sufficient or efficient attention has been paid to foster a richer conception of variable to help students overcome the difficulties they have and to help them construct a robust, flexible, and coherent conception of variable as a mathematical entity (Trigueros and Jacobs 2008), that would allow them to transit smoothly to the tertiary level of their studies.

References

Bardini, C. & Stacey, K. (2006). Students' conceptions of m and c: how to tune a linear fuction. En J. M. Moraová, H. Krátká & N. Stehlíková (Eds.) Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education (vol. 2, pp 113-120) Prague: PME.

Dorier, J. L., Robert, A., Robinet, J. & Rogalski, M. (2000). The Obstacle of Formalism in Linear Algebra. En J. L. Dorier (Ed.) On the teaching of linear algebra (pp. 83-124). Netherlands: Kluwer Academic Publishers.

Fujii, T. (2003). Probing students' understanding of variables through cognitive conflict problems: is the concept of a variable so difficult for students to understand? En N. A. Pateman, B. J. Dougherty & J. Zilliox (Eds.) *Proceedings of the Joint meeting of the PME and PME-NA* (vol. 1, pp. 49-65), Honolulu.

Gray, S. S., Loud, B. J. & Sokolowski, C. P. (2005). Undergraduates' Errors in Using and Interpreting Variables: A comparative Study. En G. M. Lloyd, M. Wilson, J. L. M. Wilkins, & S. L. Behm (Eds.) Proceedings of the 27th annual meeting of the PME-NA, Virginia Tech.

Herscovics, N. & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, 27, 59-78.

Herscovics, N. & Linchevski, L. (1996). Crossing the cognitive gap between arithmetic and algebra: operating on the unknown in the context of equations. *Educational Studies in Mathematics*, 30, 39-65.

Malisani, E. & Spagnolo, F. (2009). From arithmetical thought to algebraic thought: the role of the "variable". *Educational Studies in Mathematics*, 71(1), 19-41.

Panizza, M., Sadovsky, P. & Sessa, C. (1999). La ecuación lineal con dos variables: entre la unicidad y el infinito. *Enseñanza de las ciencias*, 17(3), 453-461.

Philipp, R. A. (1992). The many uses of algebraic variables. *The Mathematics Teacher*, 85 (7), 557-561.
Sfard, A. & Linchevski, L. (1994). The gains and the pitfalls of reification — The case of algebra, *Educational Studies in Mathematics*, 26(2-3), 191-228.
Trigueros, M. & Jacobs, S. (2008). On Developing a Rich Conception of Variable. In M. P. Carlson & C. Rasmussen (Eds.) *Making the Connection: Research and Practice in Undergraduate Mathematics*. MAA Notes #73, Mathematical Association of America, pp. 3-14.
Trigueros, M., Oktaç, A. & Manzanero, L. (2007). Understanding of Systems of Equations in Linear Algebra. Proceedings of the 5th Congress of the European Society for Research in Mathematics Education, pp. 2359-2368. Larnaka, Chipre.
Trigueros,M. & Ursini, S. (2008). Structure Sense and the Use of Variable. En O. Figueras, J.L. Cortina, S. Alatorre, T. Rojano, A. Sepúlveda (Eds.) *Proceedings of the Joint Meeting PME 32 and PME-NA XXX* (vol. 4, pp. 337-344). Morelia, México: PME & PME-NA.
Ursini, S. & Trigueros, M. (1997). Understanding of Different Uses of Variable: A Study with Starting College Students. Proceedings of the XXI PME International Conference, vol. 4, pp.254-261.
Ursini, S. & Trigueros, M. (1999). Does the understanding of variable evolve through schooling? *Proceedings of the XIII PME Conference*, vol. 4, pp. 273-280.
Ursini, S. & Trigueros, M. (2001). A model for the uses of variable in elementary algebra. *Proceedings of the XXV International Conference for the Psychology of Mathematics Education*, vol. 4, pp. 361-369.
Wagner, S. (1983). What are these things called variables? *Mathematics Teacher*, 76 (7), pp. 474-479.

20

PREPARING STUDENTS FOR CALCULUS

April Brown Judd
Northern Arizona University
Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ 86011-5717
Phone: 928-523-3481
Fax: 928-523-5847
e-mail: april.judd@nau.edu

Terry Crites
Northern Arizona University
Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ 86011-5717
Phone: 928-523-3481
Fax: 928-523-5847
e-mail: terry.crites@nau.edu

This quantitative study compared the implementation of a problem-based curriculum in Precalculus and a modular-style implementation of traditional curriculum in Precalculus to the historical instructional methods at a western Tier 2 public university. The goal of the study was to determine if either alternative approach improved student performance in Precalculus and better prepared students for success in a Calculus sequence. The study used quantitative data collection and analysis. Results indicate students who experienced the problem-based curriculum should be better prepared to learn Calculus but mixed results in terms of retention and success in Calculus.

Key words: Precalculus, calculus, problem-based learning

If Calculus is the gateway to higher-level mathematics, then Precalculus is the course that should prepare students to be students of Calculus. Students in first-semester mathematics courses continue to receive passing grades at low rates. In a report on factors effecting student success in first-year courses in business, mathematics, and science at a western Tier 2 public university, Benford and Gess-Newsome (2006) identify student academic under-preparedness and ineffective and inequitable instructional techniques as factors that contribute to the situation. The department of mathematics and statistics has been particularly concerned about the success rate of students enrolled in Calculus. Anecdotal data indicated faculty felt students entering the Calculus sequence were under-prepared. Students did not have a deep understanding of the concept of function, a "central underlying concept in calculus" (Vinner, 1992), and were not able to solve problems at the level expected in the Calculus sequence. Upon examining their preparation of students for first-semester Calculus, the department discovered students in Precalculus also experienced a low rate of passing grades (grades of C or higher).

Thus, as part of a university-wide initiative to improve student success in first-year courses with a high rate of non-passing grades (grades of D, F, W), the department of mathematics and statistics chose to examine two alternatives to the traditional curriculum in Precalculus. The goals of this initiative were to increase the rate of passing grades in Precalculus *and* Calculus and improve retention rates for students in higher-level mathematics. Historically, students participating in a Precalculus course experience lecture-based instruction, using a traditional textbook, with little opportunity to practice problems and engage with the content during class. In light of the report and faculty concerns, the department chose two alternative methods for teaching Precalculus that focused on offering students greater opportunity to master the Precalculus content, gain a deeper understanding of the concept of function, and improve their problem-solving skills.

For the first option, the department adapted a modular approach used at the University of Texas at El Paso. In this model, the Precalculus curriculum is split into three time periods, Modules 1, 2 and 3. Each module is 5 weeks in length. Students must pass an exam at the end of each module to continue to the next. If a student does not pass the exam at the end of a module, they may retake the current module over the next 5 weeks. If a student does not finish all three modules by the end of the 16-week semester, they may continue the sequence the following semester (including summer semesters). The advantage of this approach is that students are able to repeat material they have not mastered without the fear of earning a non-passing grade at the end of a traditional 16-week semester. That is, this approach gives students more time to remediate, if needed. The disadvantages are (1) instruction is not changed (i.e., students continue to experience traditional, lecture-based instruction) and (2) students must pay for an additional semester of Precalculus if they are not able to finish all three modules in a single semester.

The second option offered by the department was a reform-based curriculum focused on a quantitative approach to learning concepts in Precalculus (need to look up this reference) and a problem-based classroom environment. This curriculum was specifically designed to develop students' conceptual understanding of function (including trigonometric functions), problem solving abilities and skills that are foundational to Calculus. Students engaged in problem-based learning in groups on a daily basis. Lecture became the exception, rather than the rule, and students were expected to learn mathematics through investigating problem situations. The advantages to this curriculum are students (1) engage in solving problems every class period; (2) learn by "doing mathematics," and (3) use a research-based curriculum that reflects what students need to know to achieve success in Calculus. The disadvantage to this curriculum is that instructors and students are often unfamiliar with teaching and learning in a problem-based environment using group learning. Thus, establishing classroom norms may take longer than in a traditional college course.

The research questions for this study were as follows:

1. Does implementation of a problem-based curriculum or the adaptation of the modular approach improve student success in *Precalculus Mathematics* compared to traditional instructional methods?
2. Does implementation of a problem-based curriculum or the adaptation of the modular approach improve student preparation for *Calculus I* compared to traditional instructional methods?

Background.

The framework for this study combined ideas from work on the reasoning abilities and understandings students need to be successful in Calculus (e.g., Selden & Selden, 1999, Jensen, 2010). It is well documented that a complete notion of function, covariation, function composition, function inverse, quantity, exponential growth, and trigonometry are essential to learning in Precalculus and Calculus (Dubinsky & Harel, 1992; Rasmussen, 2000; Carlson et al., 2002; Engelke, Oehrtman & Carlson, 2005; Oehrtman, Carlson & Thompson, 2008; Carlson, Oehrtman & Engelke, 2010). Thompson, Hatfield, Byerley, & Carlson (2013) also identified several conceptual subareas important to the understanding of functions, including variables and variation, covariation, proportionality, rate of change, and structure. Although the population of the Thompson et al study was preservice secondary mathematics teachers, it is reasonable to assume that these concepts are critical to a solid understanding of functions in other groups of students as well. In addition, Stanley (2002) found that students who experience problem-based learning in Precalculus increased their ability to solve real world problems, identify and use appropriate resources, and take a more active role in their learning. Using these results, the research team chose a research-based curriculum for experimental group A that included a problem-based approach to learning and emphasized development of the function concept, covariational reasoning, and trigonometry. These results also informed the selection of the tool used to assess student preparation for Calculus (see Methodology).

Methodology.

This project used a quantitative approach of program evaluation across three types of course offerings available at a western Tier 2 public university during the 2010/2011 and 2011/2012 academic years. Quantitative methods were used to measure student success in Precalculus and preparation for first semester Calculus. In addition, qualitative methods were used to describe instructor teaching strategies that might interact with the data collected through quantitative methods. This inclusion of qualitative description helped the investigators identify any mediating variables attributed to instructional styles.

To answer research question 1, we measured overall student success in Precalculus using end-of-semester grades. To answer research question 2, we analyzed scores from the Precalculus Concept Assessment tool (PCA; Carlson, Oehrtman & Engelke, 2010) and pass/fail rates among students who completed Calculus I the semester following completion of Precalculus. The 25-item PCA multiple-choice test is a valid and reliable instrument that measures "the reasoning abilities and understandings central to Precalculus and foundational for beginning calculus." Eighteen items assess student understanding of the concept of function; five items assess student understanding of trigonometric functions; and four items assess student understanding of exponential functions. In addition, ten items require students to solve novel problem situations using quantitative reasoning and ideas of function, function composition, or function inverse. However, we recognize that instructional methods in Calculus I at this particular university might not align with research-based instructional practices in teaching and learning Calculus. Hence, we also compared student grades in Calculus I among students who completed the course the semester immediately following completion of Precalculus.

All students enrolled in Precalculus were required to complete the PCA instrument. However students were able to choose whether their PCA score was included in the study, and students' class grades were not based on their performance on the PCA. In the control group (traditional curriculum, primarily lecture-based instruction) and the experimental group A (the reform-based curriculum), the PCA was administered during the last week of classes for each semester. In experimental group B (the modular approach using a traditional curriculum), the PCA was administered during the last week of Module Three.

The PCA tool was administered as a pre/post-test during the first two academic semesters of the study to make sure the university's requirements for entering Precalculus was an accurate baseline among all students. There were no significant differences in pre-test scores between all three groups. Thus, the research team determined post-test scores were sufficient to compare differences in growth.

The Precalculus course at this particular university is a coordinated course. Instructors and Graduate Teaching Assistants (GTAs) who taught Precalculus through the traditional methods used a common course schedule of topics, a common text, and give similar exams. Similarly, the same text and schedule of topics was used for sections using the modular approach. In addition, these sections gave common exams. Instructors and GTAs using the problem-based curriculum also used a common schedule of topics, checked in regularly throughout the semester to discuss fidelity of implementation, and shared exams. However, autonomy was given to all instructors and GTAs in terms of the instructional strategies used in the classroom. That is, it was possible that one instructor might use more group work or active learning than another. To account for possible differences in instructional techniques within each type of curriculum, the researchers observed each instructor for ten minutes several times each semester. The observations were conducted by a single researcher, were informal in nature, and were meant only to verify the main form of instruction (e.g., primarily teacher-focused lecture, interactive lecture, group work, or individual problem-solving). A formal observational tool was not used. Observations took place at least three times each semester, at random times. The observer visited each classroom for ten minutes, at different times within a class period, to gain a snapshot of instruction throughout a class period for each instructor. If an instructor taught more than one section, observations were completed at different times for each class. For example, if Instructor Gary (pseudonym) taught a section at 8:10am and another section at 12:40pm, two observations may have been conducted during the first 20 minutes of the 8:10am class and two observations may have been conducted in the middle and/or latter part of the 12:40pm class. Data from these observations was used to generalize the type of instruction implemented in each course offering (i.e., traditional curriculum, modular approach, and problem-based curriculum).

Results.

The research questions for this study focused on student success in two areas of mathematics: Precalculus and Calculus I. In particular, the researchers were interested in understanding which type of curriculum best prepares students for success in Calculus within a particular population. Historically, Precalculus has been taught using a traditional text and lecture. The researchers were interested in whether adopting a modular-based approach, with a single semester broken up into three separate modules, or a problem-based curriculum would more students succeed in Precalculus and better prepare students for success in Calculus I. To answer these questions, we compared two types of scores: end-of-semester grades in Precalculus and Calculus I and total scores on the PCA instrument (described above). This section reports the results of data analysis. First we report on student success in Precalculus. Then we discuss preparation for Calculus I, as measured by the PCA instrument and students success in Calculus I as measured by end-of-semester grades. In addition, we also report our observations of instructional strategies used within each offering of Precalculus. Discussion of these results and suggestions for theory and practice are included in the following section.

Student Success in Precalculus

In order to determine if a problem-based curriculum or the adaptation of the modular approach improved student success in Precalculus compared to traditional instructional methods offered at this university, we compared end-of-semester grades for the 2010/2011 and 2011/2012 academic years (combined) using a *t*-test with the type of curriculum (tradition, modular or problem-based) used as the independent variable. It should be noted that this university transitioned out of the traditional, lecture-based curriculum after the Fall 2011 semester. Only the modular approach and the problem-based approach were offered in the Spring 2012 semester. Thus sample sizes were unequal for each type of curriculum. This variance in sample size was accounted for by using *t*-tests for unequal variances.

At this university, student success is defined as completing a course with a letter grade of A, B or C. A letter grade of D or F is considered failure since it does not earn a student credit toward their degree. Hence, we compared the mean pass rate for each type of curriculum according to the number of students who earned an A, B, or C in Precalculus. Over these two academic years, 83 percent of students in the Modular approach, 79 percent of students in the problem-based curriculum, and 71 percent in the traditional curriculum passed with a grade of C or higher. This suggests students who experienced the modular approach or the problem-based curriculum were more successful in Precalculus, with students experiencing the modular approach enjoying slightly higher success. The difference in mean pass rates was statistically significant between the traditional and modular approach and between the traditional and problem-based approach with *p*-values < 0.005. There was not a statistically significant difference between the mean pass rates for the modular approach and problem-based curriculum.

Table 1. Mean difference in pass rates for Precalculus

	Modular				Problem-based			
	M	*SD*	*t*	*p*	*M*	*SD*	*t*	*p*
Traditional	-0.12	0.20	-5.94	0.000*	-0.078	0.027	-2.896	0.004*
Modular					0.042	0.024	1.756	0.079

* indicates significance, with p < 0.005

Preparation for Calculus (PCA Scores)

Student scores on the PCA tool from the 2010/2011 and 2011/2012 academic years were compared using a *t*-test with the type of curriculum (tradition, modular or problem-based) as the independent variable. The average score on the PCA instrument among students who experienced the traditional lecture-based curriculum was 7.78; 8.63 among students who experienced the modular approach; and 10.41 among students who experienced the problem-based curriculum. Thus the mean PCA score for students who experienced the problem-based curriculum was greater than the mean score of students who experienced the traditional curriculum or the modular approach in both academic years. Furthermore, the difference in mean scores was statistically significant between all three curricula with *p*-values < 0.005.

Table 2. Mean difference in PCA scores

	Modular				Problem-based			
	M	*SD*	*t*	*p*	*M*	*SD*	*t*	*p*
Traditional	-0.851	0.236	-3.599	0.000*	-2.635	0.281	-9.374	0.000*
Modular					-1.784	0.280	-6.376	0.000*

* indicates significance, with p < 0.005

Student success in Calculus I

Student semester grades in Calculus I were compared for students who completed first semester Calculus the semester immediately after completing Precalculus. Grades were taken from the Spring 2011, Fall 2011, Spring 2012, and Fall 2012 semesters. We were only interested in whether experiencing a specific curriculum in Precalculus helped students pass Calculus I. Hence, we analyzed semester grades in terms of passing score (i.e., A, B or C). Scores were analyzed across the population of students satisfying the above requirement. We used independent sample *t*-tests to compare the pass rate in Calculus I between groups. The mean pass rate for each group was 61 percent for students who experienced the traditional curriculum in Precalculus, 69 percent for students who experienced the modular approach, and 65 percent for students who experienced the problem-based curriculum. Descriptive statistics show that the mean pass rate in Calculus I for students who experienced the modular and problem-based curriculum were slightly higher than the pass rate for students who experienced the traditional, lecture-based curriculum in Precalculus. However, the differences are not statistically significant at the $\alpha = 0.05$ level.

Table 3. Mean difference in pass rates for Calculus

	Modular				Problem-based			
	M	*SD*	*t*	*p*	*M*	*SD*	*t*	*p*
Traditional	-0.078	0.50	-1.576	0.116	-0.036	0.072	-0.496	0.621
Modular					0.043	0.069	0.615	0.540

Since the population sizes were so different for the control group and both experimental groups, we also took a simple random sample of 60 scores from each population (i.e., students who completed the traditional curriculum in Precalculus, the modular approach, or the reform-based curriculum) to verify the results above. Descriptive and *t*-test statistics for the simple random sample of 60 students in each group show similar results. The mean pass rate in Calculus I for students who experienced the modular and problem-based curriculum were slightly higher than the pass rate for students who experienced the traditional, lecture-based curriculum in Precalculus: 57 percent pass rate for students who experienced the traditional curriculum in Precalculus, 70 percent for students who experienced the modular approach, and 63 percent for students who experienced the problem-based approach. However, the differences are not statistically significant at the $\alpha = 0.05$ level.

Table 4. Mean difference in pass rates for Calculus (N=60)

	Modular				Problem-based			
	M	*SD*	*t*	*p*	*M*	*SD*	*t*	*p*
Traditional	-0.133	0.088	-1.517	0.132	-0.067	0.090	-0.741	0.460
Modular					0.067	0.087	0.770	0.443

Description of Instruction

The instructional strategies used by instructors and GTAs of Precalculus fell within the expected range for each course offering. Most instructors who taught sections using the traditional text and format for Precalculus at this university primarily used lecture-based teaching strategies. Instructors stood that the front of the room, writing notes on a chalkboard or whiteboard, with the backs to the class. Students sat in rows at long tables, taking notes. Students occasionally practiced problems on their own during a lecture. A few students checked their answers with a neighbor, but this practiced was neither encouraged nor discouraged by instructors. Two instructors gave students time to work on problem sets during class and encouraged students to work with a neighbor. In general, instructors in this group did not use group work or problem-based learning in their classes.

Instructors who taught sections using the modular approach used very similar teaching strategies. Although, after the first year of implementation, the coordinator for these sections began encouraging more problem-based teaching strategies with the use of group work. Some instructors tried using what is not known as the flipped classroom model, where students listen to an online lecture for homework. Then problem sets are worked on in groups during class. Although the online lecture and problem sets are still primarily traditional in nature, this approach does give students more time to work with peers on problems during class.

Instructors who taught sections of Precalculus using the problem-based curriculum received training prior to each semester on how to teach using a problem-based curriculum. These instructors also received instructional coaching throughout the semester by the course coordinator for these sections and used support materials provided by the authors of the curriculum. In general, most instructors implemented the curriculum faithfully. That is, group work was used every class period, and students learned material by investigating novel problem situations and discussing the material in class. One instructor found it difficult to move away from the lecture-based format. This instructor taught the problem-based curriculum for a single semester. GTAs who taught Precalculus using the problem-based approach needed more coaching from the course coordinator and took longer to fully implement a student-centered classroom. However, all GTAs using this curriculum displayed characteristics of group work and problem-based learning during every class period by the fourth week of the semester.

Conclusions.

The goal of this study was to investigate the effectiveness of two alternative options for teaching Precalculus. In particular, faculty at the participating university wished to improve student success in Precalculus and Calculus and increase the retention rate for students in higher-level mathematics. The options of a modular approach and a problem-based curriculum were chosen specifically to offer students more opportunities to master the Precalculus content, gain a deeper understanding of the concept of function, and improve their problem-solving skills. At first glance, the results presented above may seem somewhat mixed since the reported passing rates for each course do not reflect results from the PCA instrument, which is supposed to indicate readiness for first-semester Calculus. For example, end-of-semester grades and PCA scores seem to indicate that modular approach and problem-based students experienced higher success rates in Precalculus and were better prepared for first-semester Calculus than students who experienced the traditional lecture-based instruction. That is, both alternative offerings seemed to improve student success in Precalculus. The modular approach offered students more opportunities to master Precalculus content by allowing students to repeat one or more of the modules over several semesters. Whereas, the problem-based curriculum focused on helping students gain a deeper understanding of the concept of function and improve problem-solving skills.

What is not clear is if one of these two options is more helpful in preparing students for success in Calculus. Students who experienced the problem-based approach scored significantly higher on the PCA instrument than students who were taught using the modular approach or by traditional methods, indicating that these students should experience higher success rates in first-semester Calculus. However, students who completed Precalculus through the modular approach experienced higher success rates in both Precalculus (significantly better with p-levels < 0.005) and Calculus. Although the difference in mean pass rates for Calculus I were not statistically significant, one cannot ignore the overall pattern, which mirrored the statistically significant results for mean pass rates in Precalculus.

We conjecture that these seemingly contradictory results reflect the current teaching practices within the Calculus sequence, rather than on the measures used for this study. At this particular university, first-semester Calculus is primarily taught using traditional lecture-based methods. Therefore, it may be that students who completed Precalculus through the modular approach were more successful in first-semester Calculus because they had just experienced similar instruction in Precalculus and were given the opportunity to repeat material in this fashion, if needed. Thus, these students were better prepared culturally for the learning environment in Calculus I. Whereas, students who experienced the problem-based curriculum were better prepared for a first-semester Calculus course focused more on learning through the use novel problem situations, which is not offered at this university. It is yet to be determined if preparing students using a problem-based approach will encourage them to persevere into higher-level mathematics or if it is more important to prepare students for success in first-semester Calculus using instructional methods they will encounter in the future. It is also unclear if a combination of the two alternative offerings (i.e., a modular approach using more problem-based learning) would improve student success rates and retention into higher-level mathematics more so than either option on its own. Both of these questions are the goal of future investigations.

Connection to Theory and Practice.

Precalculus and Calculus I are staples of the curriculum of STEM degrees across the country. For many students, these courses are hurdles or barriers that delay or impede their degree progress. Furthermore, Calculus I instructors may often be disappointed in their students' knowledge of Precalculus concepts. While many colleges and universities deliver these courses in the traditional lecture format, others are experimenting with other methods, including problem-based and modular curricula. In theory, curricular decisions should be based on which curriculum is most likely to promote student success. In practice, other factors are also part of the curriculum decision-making process, such as the availability of financial, human, and physical resources that are needed to implement the curriculum.

The popularity of the traditional lecture format may be historical, but it probably requires the least resources. Generally, all that is needed is a chalkboard and a piece of chalk, or a PowerPoint presentation and a projector. On the other hand, modular-based curriculum can be logistically more difficult to schedule and staff. In addition, faculty probably need additional time to prepare for classes that use a problem-based curriculum than those that use a lecture format. How to balance providing the most effective curriculum and pedagogy with the reality of available resources will continue to be an issue that colleges and universities must face.

At this university, we have moved from a traditional lecture-based format to a modular-based curriculum with more problem-based instruction. It is uncertain whether this is a permanent change—only time will tell. What we do believe is that, for us, the traditional lecture format is the least effective of the three formats discussed here. This is supported by the data presented above that suggest students from Precalculus sections taught in the traditional lecture format are not as successful as those taught in the modular or problem-based format as measured by their Precalculus grade or subsequent success in Calculus.

The results of this study contribute to the knowledge base of best practices that are associated with the teaching and learning of Precalculus and Calculus. Although further research is needed, these results suggest that the traditional lecture format found in most university and college classrooms may not be the most effective method of instruction. Rather, students may learn best by being exposed to problem-based curricula that allow them to explore mathematical content in a way that develops their conceptual understanding of the mathematics instead of only their algorithmic knowledge of the procedures. We hope that these results will prompt Precalculus teachers (including those at our own university) to reexamine their instructional strategies and practices.

References

Alkhateeb, H., & Mji, A. (2005). Attitudes of university precalculus students toward mathematics. *Psychological Reports*, *96*(2), 257-258.

Bandura, A. (2001). Social foundations of thought and action: A social cognitive theory. *Annual Review of Psychology*, 52, 1-26.

Benford, R., & Gess-Newsome, J. (2006). Factors Affecting Student Academic Success in Gateway Courses at Northern Arizona University. *Online Submission.*

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.

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.

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, 25, 85–106.Washington, DC: Mathematical Association of America.

Engelke, N., Oehrtman, M., & Carlson, M. (2005). Composition of functions: Precalculus students' understandings. Proceedings of the 27th AnnualMeeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Roanoke: Virginia Tech.

House, J. D. (1995) The predictive relationship between academic self-concept, achievement expectancies, and grade performance in college calculus. Journal of Social Psychology, 135, 111-112.

Jensen, T. (2010). A study of the relationship between introductory calculus students' understanding of function and their understanding of limit. *Dissertation Abstracts International Section A*, *70*,

Lester, F. K., Garofalo, J. & Kroll, D. L. (1989) Self-confidence, interest, beliefs and metacognition: key influences on problem-solving behavior. In D. B. McLeod & V. M. Adams (Eds.), Affect and Mathematical problem solving: a new perspective. Near York: Springer-Verlag. Pp. 75-88.

Oehrtman, M., Carlson, M., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. Carlson & C. Rasmussen (Eds.), Making the connection: Research and practice in undergraduate mathematics, MAA Notes Volume, 73, 27–41. Washington, DC: Mathematical Association of America.

Piper, Bradley D. (2008). Attitudes, Confidence, and Achievement of High-Ability Fifth Grade Math Students. In partial fulfillment of the MA Degree, Department of Teaching, Learning, and Teacher Education, University of Nebraska. Retrieved on August 17, 2010 from DigitalCommons@University of Nebraska – Lincoln, http://digitalcommons.unl.edu/mathmidsummative/29.

Randhawa, B.S., Beamer, J. E. & Lundberg, I.(1993). Role of mathematics self-efficacy in the structure model of mathematics achievement. Journal of Educational Psychology, 85, 41-48.

Rasmussen, C. (2000). New directions in differential equations: A framework for interpreting students' understandings and difficulties. Journal of Mathematical Behavior, 20, 55–87.

Selden, A., Selden, J., Hauk, S., Mason, A., & Tennessee Technological University, D.

(1999). Do Calculus Students Eventually Learn to Solve Non-Routine Problems? Technical Report. No. 1999-5. *Online Submission.*

Stanley, S. S. (2002). Revitalizing Precalculus with Problem-Based Learning. *Journal Of General Education*, *51*(4), 306-315.

Thompson P., Hatfield, N.,Byerley, C. & Carlson, M. (2013). Under the Radar: Foundational Meanings that Secondary Mathematics Teachers Need, Do Not Have, but Colleges Assume. Proceedings of the Sixteenth Annual Conference on Research in Undergraduate Mathematics Education. Paper presented at the Sixteenth Annual Conference on Research in Undergraduate Mathematics Education, Denver, Colorado, 21 -23 February.

Vinner, S. (1992). The function concept as a prototype for problems in mathematics learning. In G. Harel & E. Dubinsky (Eds.), The concept of function: Aspects of epistemology and pedagogy. MAA Notes, 25, 195–213.

COOPERATIVE LEARNING AND TRAVERSING THE CONTINUUM OF PROOF EXPERTISE: PRELIMINARY RESULTS

Martha Byrne
University of New Mexico

This paper describes preliminary results from a larger study aimed at examining the effects of working in cooperative groups on acquisition and development of proof skills. Particular attention will be paid to the varying tendencies of students to switch proof methods (direct, induction, contradiction, etc) based on their level of proof expertise. Namely, as students progress from novice to expert provers, they tend to change proof methods more frequently until they reach the final stages of development (Hart 1994).

Key words: Transition to Proof, Proof Writing Expertise, Proof Methods

Introduction

Although proof is essential to studying mathematics, much research in the past two decades shows that students struggle with constructing and validating proofs (Almeida, 2000; Harel & Sowder, 1998; Levine & Shanfelder, 2000, Moore, 1994; Selden & Selden, 2003a, 2003b; Weber, 2001; Weber, 2003). Several innovative course structures have been introduced for so-called bridge courses (Almeida, 2003; Bakó, 2002; Grassl & Mingus, 2004), but little dedicated research has been done on the effectiveness of such courses. However, some common themes have emerged about the necessity for and efficacy of active learning strategies, and there is a general trend away from lecture and toward more student-centered models. In particular, this can be seen within the Modified Moore Method community (McLoughlin, 2010).

Cooperative learning (CL) is one such model. “CL may be defined as a structured, systematic instructional strategy in which small groups work together to produce a common product” (Cooper, 1990). There are six specific features that, when combined, distinguish CL from other active and collaborative learning strategies: positive interdependence, individual accountability, appropriate grouping, student interaction, attention to social skills, and teacher as facilitator. While the efficacy of CL has been researched (Johnson & Johnson, 1991), the majority of this research has been undertaken with precollegiate populations.

Studies done on CL and active learning in the context of physics instruction (Deslauriers, et al, 2011; Heller & Hollabaugh, 1992; Heller, et al., 1992) give hope that CL could be effective in helping students acquire and develop their proof skills. This paper looks at some of the preliminary results of a study exploring the relationship between CL and proof-skill development. Specifically, the study was designed to examine how working in a CL seminar environment affected 1) student attitudes about proof, 2) student ability to construct proofs, and 3) student ability to validate student-generated arguments. The second of these will be addressed in this paper.

Hart (1994) compared expert and novice proof writers through use of a proof test he developed for the study. He categorized 29 undergraduate math majors by their level of proving expertise using three specific tasks from the test, and rated the students in one of four levels: Level 0: pre-understanding, Level 1: syntactic understanding, Level 2: concrete semantic understanding, Level 3: abstract semantic understanding (p. 56). He then examined the students’ individual proof production processes, noting similarities that arose among students at the same level of understanding.

The mathematical context of Hart’s study, abstract algebra, differed from that of the study addressed in this paper, but he reported some generally applicable findings. He noted that novice provers do not perform like expert provers right away; there are several stages of expertise that must be traversed, even though the progression through these stages often does not proceed at a steady rate. In particular, he noted that expert provers switch proof methods less often than the most novice provers, but that the tendency to change plans increased between levels at all but the final step (pp. 59-60). This study examined whether students working to develop proofs in a cooperative learning environment would show measurable increases in their proving skills as measured by their success in constructing proofs and their tendency to change plans.

Methods

This study looked at the change in proof construction skills of undergraduate students at a large, public university who had declared majors or minors in mathematics or mathematics education. Participants were given pre- and post- assessments consisting of a proof construction task at the beginning and end of the semester (respectively). The study consisted of a treatment group and a comparison group. Initially, eight treatment students and six comparison students took the pre-assessment, but one treatment student and three comparison students dropped out of the study after the pre-assessment.

The remaining seven treatment participants (five male, two female) met collectively with the researcher for eight 90-minute seminar sessions over the course of the semester between pre- and post- assessments. The students were assigned into groups based on gender, previous proof experience, and performance on the pre-assessment. During the seminar sessions, the students worked in their groups on proof-based problem sets designed by the researcher. Each seminar participant but one was enrolled in at least one proof-based course at the university. The other three participants who remained in the study (two male, one female) were also enrolled in at least one proof-based course at the university. None of the courses the students were enrolled in utilized any structured group work as a part of the course. Between assessments, these students had no interactions with the researcher, and while they participated in their proof-based courses engaged in no group work activities. For this paper, the students will be referred to as seminar students and comparison students respectively.

The pre- and post- assessments consisted of three proof prompts in basic number theory. Each was presented as a true theorem to the participants. Elementary number theory prompts were chosen so the necessary concepts would be accessible to all of the subjects regardless of prior background. Each prompt was selected to test additional proof construction skills the researcher believes are context-independent proof skills (see Table 1). The assessments of the seminar students were conducted in the presence of the researcher, and a think-aloud protocol was employed as they attempted to construct the proofs. The assessments of the comparison students were conducted in a group setting with each subject working independently and silently. Pre-assessments were taken during the first four weeks of the spring 2012 semester, and post-assessments were taken during the final week and the first week following the spring 2012 semester. At least 11 weeks passed between pre- and post- assessments for all subjects. Pre- and post-assessments for both sets consisted of a mathematical background questionnaire, proof attitudes survey, proof construction task, and proof validation task. The seminar students had an additional research experience interview during the post-assessments.

During the seminar sessions between assessments, the seminar students worked on problem sets in their assigned cooperative groups. The cooperative groups were consistent throughout the study and were formed so that the female students were not outnumbered in the groups and so that they were heterogeneous based on skill level as demonstrated on the pre-assessment. The members of each group spent a few minutes at the beginning of each session getting to know each other and 5-10 minutes at the end of each session doing group processing exercises. Both of these exercises facilitated the development of the social skills necessary for effective cooperative work, and the rotating roles (manager, explainer, skeptic, presenter) the students assumed each session assured their personal accountability and positive interdependence. After a brief introduction each session, the students worked with each other and the researcher functioned solely as a facilitator, encouraging the student-to-student interactions.

The problem sets dealt with function concepts, primarily injectivity and surjectivity, and the seminar group did not work with number theoretic concepts. This was done so that any changes from pre- to post-assessment would reflect changes in the subjects' proving skills independent of mathematical context.

Assessment Item	Proof Skill(s) Tested
1. Prove: If m^2 is odd, then m is odd.	A. Use of indirect proof methods. B. Avoidance of a more appealing but logically inequivalent converse argument.
2. Prove: If n is a natural number, then n^3-n is divisible by 6.	A. Ability to identify pertinent subclaims and construct subarguments (divisibility by 2 and 3).
3. A *triangular number* is defined as a natural number that can be written as the sum of consecutive integers, starting with 1. Prove: A number, n, is triangular if and only if $8n+1$ is a perfect square. (You may use the fact that $1+2+...+k = \frac{k(k+1)}{2}$.)	A. Use of the specifics of a definition to form a basis for a proof. B. Ability to identify the logical implications of "if and only if" statements. C. Use of previously established results (to prove $8n+1$ a perfect square implies that n is triangular, the result of item one needs to be applied).

Table 1. Assessment Items

Data Analysis

Data from Seminar Students

The researcher met with each of the seminar students individually to conduct the pre- and post-assessments. All of these assessments were videotaped as were the seminar sessions, and all videos were transcribed. To analyze the data gathered from the seminar students, the researcher used an open coding scheme to examine the transcripts and written work and interviews for evidence of the specific proof skills tested by the items on the pre- and post-assessments (see Table 1).

The written work was then examined again and each proof attempt was coded using Boyle's (2012) coding scheme shown in Table 2. In this scheme, each argument is given one of four broad categories ranging from incoherent or absent arguments (A0) to valid proofs (A4). For unsuccessful proof (categories A0, A1, A2, A3), specific errors and flaws were noted and encoded. For successful arguments (categories A3 and A4), three specific aspects of clarity were addressed. It is worthwhile to note that attempts that fell into category A3 represented valid logical arguments but not rigorous proofs. When subclaims or subarguments were addressed by the students, the subarguments were coded using the same scheme. See Figure 1 for an example of a student's argument with coding.

During the examination of the written work and transcripts, the researcher found that during the post-assessments, the seminar students seemed to be much more willing to attempt different types of proof and change proving methods when stuck than they had on the pre-assessment. Because of this observation, the researcher went back to the transcripts and written work again and counted the number of times each student changed plans for each proof attempt. Any time a student started working on a type of proof (direct, contradiction, contrapositive, or induction) or came back to an argument that had previously been abandoned, the researcher counted the action as a method switch.

Argument Codes	Code Details	Code Evidence
Incoherent or not addressing the stated problem (A0)	1. Solution shows a misunderstanding of the mathematical content. 2. Ignores the question completely.	• List A0 and either 1 or 2
Empirical (example based) (A1)	1. Examples are used to find a pattern, but a generalization is not reached. 2. Only examples are generated as a complete solution.	• List A1 and either 1 or 2
Unsuccessful attempt at a general argument (A2)	1. There is a major mathematical error 2. Illogical reasoning; several holes and or errors exist causing an unclear or inaccurate argument. 3. Reaches a generalization from examples, but does not justify why it is true for all cases. 4. Solution fails to covers all cases. 5. Solution is incomplete. Argument stops short of generalizing the stated claim.	• List A2 and match the bulleted number (1-5) in the middle column with the work in the solution.
Valid argument but not a proof (A3)	1. The solution assumes claims, in other words the solution exhibits a leap of faith before reaching a conclusion 2. The solution assumes a conjecture or lists a non-mathematical statement as a conjecture.	• List A3 and either 1 or 2 & address each of the points below **
Proof (A4)		• List A4 and address each of the three clear and convincing points below. **
** for use with A3 and A4. (+/-) The flow of the argument is coherent since it is supported with a combination of pictures, diagrams, symbols, or language to help the reader make sense of the author's thinking. (+/-) There are no irrelevant or distracting points. Variables and definitions are clearly defined. (+/-) The conclusion is clearly stated.		

Table 2. Proof Codes

When it became apparent that there was a general increase in switching methods from pre- to post- assessment for the seminar students, the researcher used an open coding scheme to examine the seminar transcripts for evidence that the seminar supported this tendency. Finally, the researcher examined the transcripts of the seminar students' research experience interviews and looked for trends in what the students believed had impacted their learning and habits over the course of the study.

Data from Comparison Students

The comparison students' assessments were conducted all at once, and while the researcher was present, she did not converse with the students about their work or experiences. Their written proof attempts were coded using Boyle's scheme (see Table 2) and examined for evidence of the tested proof skills. There was no evidence in any of the written attempts that any of the comparison students had switched methods while working. However, this does not mean that they did not think about doing so.

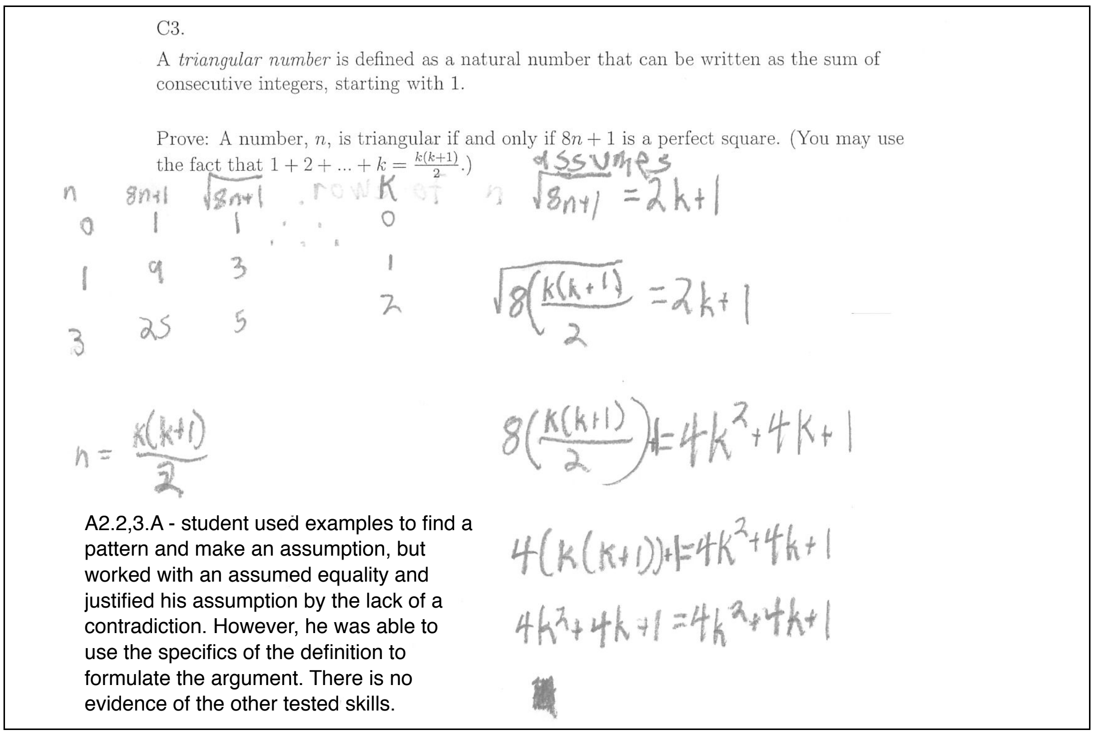


Figure 1. Ivan's Pre-Assessment Task 3

Preliminary Findings

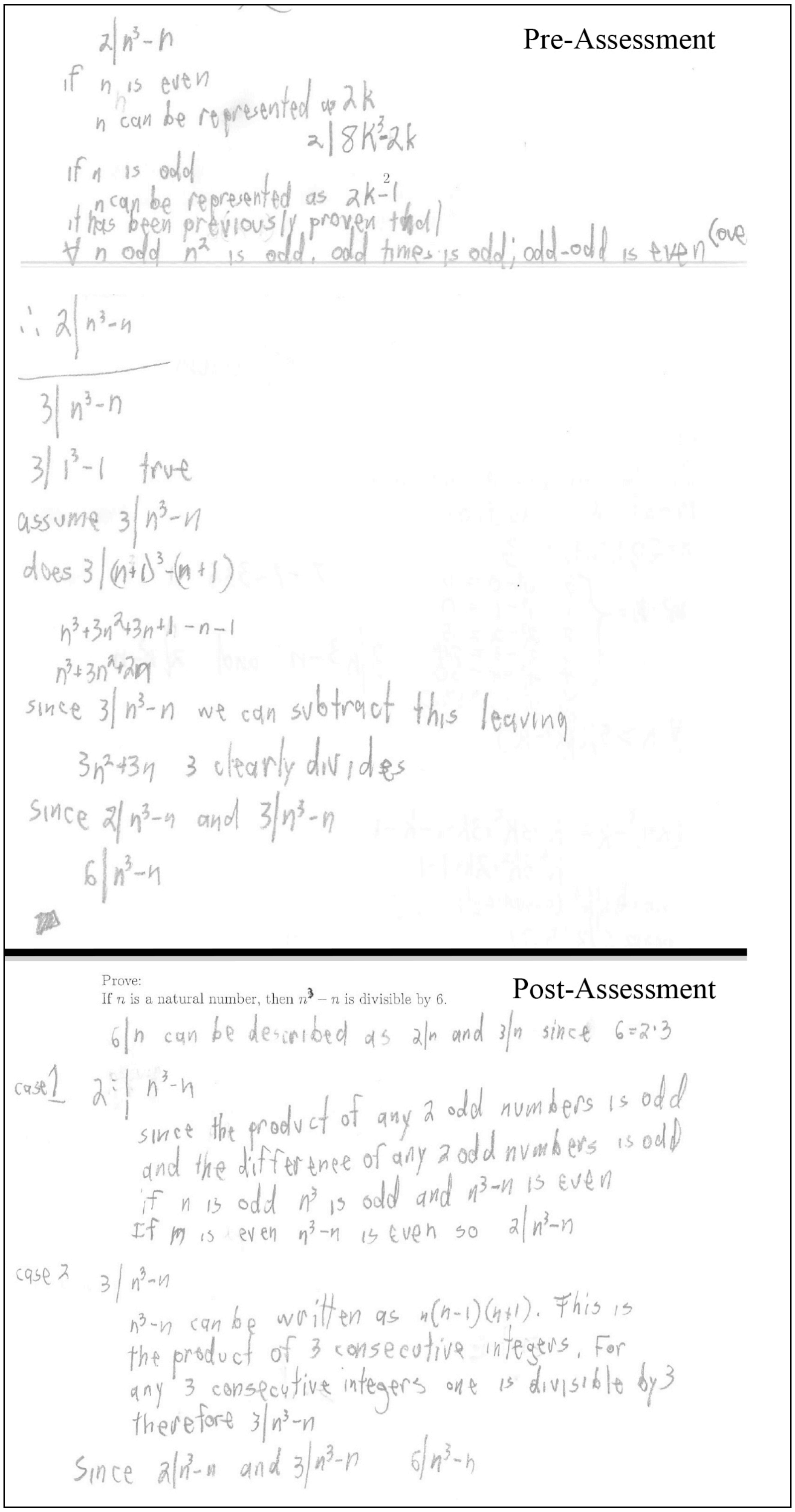
Results for Seminar Students

Of the seven seminar students, six showed distinct improvement from pre-assessment to post-assessment (see Table 3). Improvement was defined as increased argument code on the post-assessment (e.g. any argument rated an A2 was seen as worse than any argument rated A3 or A4) and/or post-assessment evidence of a tested proof skill that was not apparent on the pre-assessment. For example, Ursula improved on both item 1 and item 2 as shown in Table 3 despite the fact that item 2 was rated an A2 on both assessments because she demonstrated an ability to construct subclaims and subarguments on the post-assessment, but not on the pre-assessment. Of the eighteen composition item comparisons for these students (e.g. Ivan's pre-assessment task 2 was compared to Ivan's post-assessment task 2), there was improvement on twelve, stasis on five, and regression on only one. The regression was mild, and it is the researcher's belief that the student miswrote what he was thinking (see Figure 2). Regardless,

Student	PRE			POST		
	Item	Subclaims	CODE	Item	Subclaims	CODE
Bill	**1**		**A2.1.A,B**	**1**		**A4.+++.A,B**
	2		**A2.5.A**	**2**		**A2.5.A**
		by 2	A4.+++		by 2	A4.+++
		by 3	A0.2		by 3	A0.2
	3		**A2.5.A,B**	**3**		**A3.2+++.A,B**
		T -> S	A0.2		T -> S	A4.+++
		S -> T	A2.5		S -> T	A3.2.++
Ingrid	**1**		**A2.1,2.B**	**1**		**A4.+-+.A.B**
	2		**A2.1,5.A**	**2**		**A2.2,5.A**
		by 2	A2.3.5		by 2	
		by 3	A0.1,2		by 3	
	3		**A2.5.A,B**	**3**		**A3.2.++-.A,B**
		T -> S	A0.2		T -> S	A4.-+-
		S -> T	A0.2		S -> T	A3.2.++-
Ivan	**1**		**A2.2.N**	**1**		**A2.2.N**
	2		**A4.+++.A**	**2**		**A2.1**
		by 2	A4.+++		by 2	A2.1
		by 3	A4.+++		by 3	A4.+++
	3		**A2.2,3.A**	**3**		**A2.5.A,B**
		T -> S	A0.2		T -> S	A4.-++
		S -> T	A0.2		S -> T	A0.2
Nathan	**1**		**A4.---.A,B**	**1**		**A4.-++.A,B**
	2		**A2.5.N**	**2**		**A2.2,4.N**
	3		**A2.5.A,B**	**3**		**A4.-+-.A,B,C**
		T -> S	A4.---		T -> S	A4.---
		S -> T	A2.5		S -> T	A4.---
Omar	**1**		**A2.2.N**	**1**		**A4.+++.A,B**
	2		**A0.1**	**2**		**A2.2,5.N**
	3		**A0.2**	**3**		**A2.5.A**
		T -> S	A0.2		T -> S	A2.5
		S -> T	A0.2		S -> T	A0.2
Ursula	**1**		**A2.2.A,B**	**1**		**A4.-+-.A,B**
	2		**A2.1,5.N**	**2**		**A2.5.A**
					by 2	A0.2
					by 3	A0.2
	3		**A2.1,2.A,B**	**3**		**A2.2,5.A,B**
		T -> S	A2.1,2		T -> S	A2.1
		S -> T	A0.2		S -> T	A0.2
Zach	**1**		**A4.++-.A,B**	**1**		**A2.2.B**
	2		**A2.1,5.A**	**2**		**A2.5.A**
					by 2	A0.2
					by 3	A0.2
	3		**A2.5.N**	**3**		**A2.2,3.N**
		T -> S	A0.2		T -> S	A0.2
		S -> T	A0.2		S -> T	A0.2

Table 3. Seminar Students' Performance

each of those six students showed significant improvement on at least one task. The seventh student, Zach, regressed on one of the composition items and showed no improvement on the other two. It should be noted that he spent less time on each item than he had during the pre-assessment and expressed his frustration with the fact that the seminar had not covered the concepts involved in the assessment.



Pre-Assessment

$2 \mid n^3-n$

if n is even

n can be represented w $2k$

$2 \mid 8k^3-2k$

if n is odd

n can be represented as $2k-1$

it has been previously proven that

$\forall\, n$ odd n^2 is odd, odd times is odd; odd-odd is even

$\therefore 2 \mid n^3-n$

$3 \mid n^3-n$

$3 \mid 1^3-1$ true

assume $3 \mid n^3-n$

does $3 \mid (n+1)^3-(n+1)$

$n^3+3n^2+3n+1-n-1$

n^3+3n^2+2n

since $3 \mid n^3-n$ we can subtract this leaving

$3n^2+3n$ 3 clearly divides

Since $2 \mid n^3-n$ and $3 \mid n^3-n$

$6 \mid n^3-n$

Post-Assessment

Prove:
If n is a natural number, then $n^3 - n$ is divisible by 6.

$6 \mid n$ can be described as $2 \mid n$ and $3 \mid n$ since $6=2\cdot3$

case 1 $2 \mid n^3-n$

since the product of any 2 odd numbers is odd and the difference of any 2 odd numbers is odd

if n is odd n^3 is odd and n^3-n is even

If n is even n^3-n is even so $2 \mid n^3-n$

case 2 $3 \mid n^3-n$

n^3-n can be written as $n(n-1)(n+1)$. This is the product of 3 consecutive integers. For any 3 consecutive integers one is divisible by 3 therefore $3 \mid n^3-n$

Since $2 \mid n^3-n$ and $3 \mid n^3-n$ $6 \mid n^3-n$

Figure 2. Ivan's Pre- and Post-Assessment Work for Item 2

Most of the seminar students other than Zach changed proof methods more frequently on the post-assessment than they had on the pre-assessment. The students who had the greatest change were those who had the weakest performances on the pre-assessment. In fact, Bill was the only student who changed proof methods less frequently on the post-assessment than on the pre assessment, and he was the strongest student on the pre-assessment (see Table 4).

Seminar Student	Total Number of Switches	
	Pre-Assessment	Post-Assessment
Omar	0	1
Ursula	3	6
Ingrid	0	6
Ivan	3	3
Zach	4	0
Nathan	1	2
Bill	5	4

Table 4. Seminar Student Switching Tendency on Assessments (students ranked low to high by pre-assessment performance)

The participants naturally fell into two categories on the pre-assessment: a low performing group consisting of Omar, Ursula, Ingrid, and Ivan, and a high group consisting of Zach, Nathan and Bill. All members of the high group except Zach had items on the post-assessment for which they changed plans less frequently but performed as well or better. However, all members of the low group changed plans at least as many times on every item on the post-assessment as they had on the pre-assessment (See Tables 5a and 5b).

BILL	Description of Performance	Code	# of Switches
Item 1 - pre	Defined m^2 as odd by setting it equal to $2x^2+1$, which is a mathematical error. Produced an otherwise valid proof by contradiction.	A2.1.A,B	3
Item 1 - post	Produced a valid proof by contrapositive	A4.+++.A,B	1
Item 2 - pre	Produced a proof that n^3-n is even, and recognized he was missing that 3| n^3-n.	A2.5.A	3
Item 2 - post	Produced a proof that n^3-n is even, and recognized he was missing that 3| n^3-n.	A2.5.A	1
Item 3 - pre	Manipulated the equation $8n+1=x^2$, but the manipulations were unproductive.	A2.5.A,B	0
Item 3 - post	Produced a proof of both directions, but was missing the justification that $8n$+1 is necessarily odd, so if it is a perfect square, then it is the square of an odd number.	A3.2.+++.A,B	2

Table 5a. Bill's Performance and Switching on Individual Items

URSULA			
Item 1 - pre	Produced an empirical contradiction argument with use of a single example, m=2.	A2.2.A,B	1
Item 1 - post	Produced a valid proof by contrapositive.	A4.-+-.A,B	2
Item 2 - pre	Did not identify subgoals, attempted a proof by induction, but could not get to conclusion even though the work was error-free.	A2.1,5.N	1
Item 2 - post	Identified subgoals, attempted a proof by contradiction again and successfully proved that $3\mid n^3\text{-}n$. Had the work to get $2\mid n^3\text{-}n$, but did not recognize that $k^2\text{-}k$ is necessarily even.	A2.5.A	3
Item 3 - pre	Correctly stated givens and goals, attempted to use $1+2+...+k=\frac{k(k+1)}{2}$, but set up $8(\frac{k(k+1)}{2})+1=triangular$. Made no progress from there.	A2.1,1.A,B	1
Item 3 - post	Proved that n triangular implies $8n+1$ is a perfect square. Made no progress on reverse direction.	A2.2,5.A,B	1

Table 5b. Ursula's Performance and Switching on Individual Items

Results for Comparison Students

Of the three comparison students, two demonstrated evidence of a tested proof skill on one item of the post assessment that had not been apparent on the pre-assessment. However, none of the comparison students achieved a higher argument code on the post-assessment than had been achieved on the comparable pre-assessment item. All told, of the nine item comparisons, improvement was demonstrated only on the two items mentioned above, while there was stasis on six and regression on one (see Table 6).

Student	PRE			POST		
	Item	Subclaims	CODE	Item	Subclaims	CODE
0296	**1**		**A4.+++.A,B**	**1**		**A4.++-.A,B**
	2		**A2.5.N**	**2**		**A2.5.N**
	3		**A0.2.N**	**3**		**A0.2.A**
4586	**1**		**A4.+++.A**	**1**		**A4.+++.A**
	2		**A3.1.N**	**2**		**A3.1.N**
	3		**A2.5.A,B**	**3**		**A2.5.A,B**
		T -> S	A4.+++		T -> S	A4.+++
		S -> T	A2.5		S -> T	A2.5
6772	**1**		**A2.2.N**	**1**		**A2.2.A**
	2		**A2.1.N**	**2**		**A2.1,5.N**
	3		**A3.2+++.A,B**	**3**		**A2.5.A,B**
		T -> S	A4.+++		T -> S	A4.+++
		S -> T	A3.2+++		S -> T	A0.2

Table 6. Comparison Students' Performance

In contrast to Ivan's slight regression, the regression demonstrated by comparison student 6772 was more significant. On the third item of the pre-assessment, this student produced a nearly complete argument. A flawless argument of one direction was present, and the argument for the second direction included a valid, but unjustified assumption. However, on the post assessment the first argument was still valid, but the second argument was completely absent. There is no evidence that any of the comparison students switched proof methods at any point during either of the assessments.

Discussion

Limitations of the Study

This study involved a very small number of volunteer participants. Both because of the small sample size and because the participants volunteered and were allowed to choose their participation level (seminar or comparison), no definitive conclusions can be drawn about what caused the seminar students' improvement over the comparison students' relative stasis. However, there is a distinct difference in the progress of the two groups that is consistent with the researcher's hypothesis that group work could contribute to the acquisition and development of proof skills.

Additionally, the fact that the assessments were not conducted in the same environments for the two sets of participants is problematic because most of the evidence that the seminar students were switching proof methods was found in the interview transcripts and not in the students' written work. The absence of any evidence of comparison students switching methods is potentially a consequence of the assessment format and it may not be the case that they did not switch methods on the assessments.

Conclusions

During the seminar sessions, the cooperative groups worked collectively on problem sets created by the researcher. These problem sets often asked the students to outline the arguments that would be needed to prove a given claim directly, by contradiction, and by contrapositive. There were also several discussions with all the seminar students and the researcher about when to employ the differing proof strategies and how to tell whether or not such an outlined strategy might be effective or ineffective. These discussions and activities could reasonably be responsible for the some of the increased tendency of the seminar students to change plans. The researcher hypothesizes that this increased flexibility helped enable the seminar students to be more successful on their post-assessments and can explain why the same increase would not have been present for the comparison students.

While the results from the seminar participants are consistent with Hart's findings, it is certainly not the case that expert provers always exhibit a low tendency to switch proof methods. Especially when working in an unfamiliar area, even research mathematicians will sometimes switch proof methods multiple times while trying to prove a single theorem. This researcher hypothesizes that both comfort level with the mathematical context and proving expertise are needed for a prover to be able to consistently choose a productive proof method at the beginning of a proof attempt. However, for a novice prover, unfamiliarity with alternative proof methods and/or an unwillingness to change proof methods when at an impasse generally results in an inability to consistently prove theorems. In this study, we saw that increased flexibility was generally paired with improved performance, although that flexibility itself doesn't necessarily indicate a students' position along a proving expertise continuum.

Implications for Future Research

Because of the limitations mentioned above, no strong conclusions about causation can be drawn from this study, but the contrast between the seminar and comparison students' improvement, is consistent with the researcher's hypothesis and supports the undertaking of a larger, blind study in the future to test the contribution of the group work and the specific emphasis on alternative proof methods to the students' proof skill development.

Additionally, there is some debate in the undergraduate math education research community about the existence of content-independent proving skills, but this study is consistent with the hypothesis that such skills exist. More analysis on these data and additional study is merited to explore that question further.

References

Almeida, D. (2000). A survey of mathematics undergraduates' interaction with proof: some implications for mathematics education. *International Journal of Mathematical Education in Science and Technology*. 31, #6, 869-890.

Almeida, D. (2003). Engendering proof attitudes: can the genesis of mathematical knowledge teach us anything?. *International Journal of Mathematical Education in Science and Technology,* 34(4), 479-488.

Bakó, M. (2002). Why we need to teach logic and how can we teach it? *International journal for Mathematics Teaching and Learning*. Retrieved July 24, 2012 from www.cimt.plymouth.ac.uk/journal/bakom.pdf

Boyle, J. (2012). A Study of Prospective Secondary Mathematics Teachers' Evolving Understanding of Reasoning-and-Proving. (EdD), University of Pittsburgh.

Cooper, J. (1990). What is Cooperative Learning? In Cooper, J., Robinson, P., & Ball, D. (Eds.), *Small Group Instruction in Higher Education: Lessons from the Past, Visions of the Future*. (pp. 3-5). Stillwater, OK: New Forums Press, Inc.

Deslauriers, L., Schelew, E., & Wieman,C. (2011). Improved Learning in a Large-Enrollment Physics Class. *Science*, 13 May 2011: Vol. 332 no. 6031 pp. 862-864

Grassl, R. & Mingus, T. T. Y. (2004). Team teaching and cooperative groups in abstract algebra: nurturing a new generation of confident mathematics teachers. *International Journal of Mathematical Education in Science and Technology.* 38(5), 581-597.

Harel, G., & Sowder, L. (1998). Students' proof schemes. *Research on Collegiate Mathematics Education, Vol. III*. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), AMS, 234-283.

Hart, E. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. In Kaput, J. & Dubinsky, E. (Eds.), *Research issues in undergraduate mathematics learning: Preliminary analyses and results*, (pp. 49-63). Washington, DC: The Mathematical Association of America.

Heller, P., & Hollabaugh, M. (1992). Teaching Problem Solving Through Cooperative Grouping Part 2: Designing Problems and Structuring Groups. *American Journal Physics*, 60(7), 637-644.

Heller, P., Keith, R. & Anderson, S. (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics*, 60(7), 627-636.

Johnson, D. W. & Johnson, R. T. (1990). Social Skills for Successful Group Work. *Educational Leadership*, 48(7), 29-33.

Johnson, D.W., & Johnson, R.T. (1991). *Cooperation in the classroom.* Edina, MN: Interaction Book Company.

Levine, A. & Shanfelder, B. (2000) The Transition to Advanced Mathematics. *PRIMUS*, 10(2), 97–110.

McLoughlin, P. (2010, August) *Aspects of a Neoteric Approach to Advance Students' Ability to Conjecture, Prove, or Disprove*. Paper presented at the annual meeting of the The Mathematical Association of America MathFest, Omni William Penn, Pittsburgh, PA

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

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.

Selden, A. & Selden, J. (2003). Errors and misconceptions in college level theorem proving. In Novak, J. D. (Ed.) *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Vol. III*, Cornell University, July 1987, 457-470.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, 48(1), 101-119.

Weber, K. (2003). A procedural route toward understanding the concept of proof. *Proc. PME* 27, Vol. 4, pp. 395-401.

PROVIDING CALCULUS STUDENTS WITH OPPORTUNITIES TO ENGAGE IN THEORETICAL THINKING

Dalia Challita Nadia Hardy
Department of Mathematics and Statistics
Concordia University, Montreal, Canada

In this research, we provide empirical evidence that students can be engaged in theoretical thinking in a university introductory level Calculus course, despite the institutional constraints that often surround and pervade these courses. Students enrolled in a Calculus course were presented with optional tasks intended to engage them in theoretical thinking. We analyze the results from the perspective of Sierpinska et al.'s (2002) model for theoretical thinking; our analysis shows that students who participated in these optional tasks often engaged in theoretical thinking. We discuss these findings in the context of previous research in the teaching and learning of university introductory (and remedial) level mathematics and of the role that Calculus courses play in the mathematics education of undergraduate students.

Keywords: Theoretical thinking, Calculus, Quizzes, Institutional constraints

Research carried out in the last two decades shows that university mathematics introductory courses do not involve students in theoretical thinking (TT) (e.g., Hardy & Challita, 2012; Lithner, 2003; Selden et al., 1999). In particular, the types of problems that students are typically presented with in these courses do not invite TT (e.g., Boesen et al., 2010; Hardy, 2009b; Lithner, 2003; Raman, 2004; Selden et al., 1999) and students' discourses about the validity and pertinence of a problem-solving technique are of social and didactic nature rather than of mathematical nature (Hardy, 2009a). Researchers attribute the absence of a TT component in these introductory courses to the teaching style (e.g., Hardy & Sierpinska, 2011; Sierpinska et al., 2002) and existing institutional constraints (e.g., Barbé et al., 2005; Hardy, 2009a). As mathematics educators and researchers in mathematics education, we presume that teaching a subject matter involves offering students situations in which they can display and practice behaviors characteristic of experts in the given subject matter. In mathematics, these behaviors involve, among other things, the use of particular problem solving strategies (Schoenfeld, 1987), TT (Sierpinska et al., 2002), and a preoccupation for theoretical consistency and validation (Hardy, 2009a). In this paper, we provide empirical evidence that it is indeed possible to engage Calculus students in TT, despite the institutional constraints that often surround and pervade these courses in university settings.

Institutional Constraints

North-American universities typically offer introductory Calculus courses as pre-requisites for programs as varied as Natural Sciences, Business, Engineering, Computer Science, Health Science, etc. These courses are often offered in a multi-section format, taught by different instructors but designed by a single course examiner/coordinator. The course examiner writes the outline of the course, the weekly assignments, and the midterm and final exams. In addition to specifying the topics to be taught each week, the outline includes a list of "recommended" exercises from a common-assigned textbook which are indicative of the types of problems that will appear on the weekly assignments and midterm and final exams. The problems on these assessments can be described as "routine" in the sense of (Lithner, 2003; and Selden et al., 1999); dealing with these tasks seldom, if at all, invites or requires thinking theoretically. Instructors of these courses face several institutional constraints: a fixed outline (i.e., they cannot change it), with a fixed order for delivering the content, a pre-chosen set of exercises, assessments that they cannot modify; plus constraints associated to classroom time and class-size. Our goal is to explore whether in a course and setting such as the one described here, and despite the mentioned constraints imposed on the instructor, there is place for tasks that invite TT – in the sense of what format could these tasks take, considering the institutional constraints, and in the sense of whether students presented with such tasks would actually engage in them and display the use of TT.

Theoretical Perspective

To characterize TT and analyze students' uses of it, we consider the model of TT constructed by Sierpinska et al. (2002). This model is based on Vygotsky's characterization of scientific concepts as formed in the mind on the basis of concretization from general statements, as opposed to spontaneous concepts that are formed on the basis of generalization and verbalization from concrete experience. In Sierpinska et al.'s (2002) model, the main postulated categories of theoretical thinking are *reflective*, *systemic*, and *analytic* thinking (a shorthand term for *analytic approach to signs*), with systemic and analytic thinking broken down further into "features" of TT. (See Table 1)

Category of TT **• Feature of TT**	**Description**
Reflective	Thinking for the sake of thinking.
Systemic	Thinking about systems of concepts, where the meaning of a concept is established based on its relations with other concepts and not with things or events.
• Definitional	Stabilizing the meanings of concepts by means of definitions.
• Proving	Being concerned with the internal coherence of conceptual systems.
• Hypothetical	Being aware of the conditional character of statements; hypothetical thinking seeks to uncover implicit assumptions and study all logically conceivable cases.
Analytic	Having an analytical approach to signs
• Linguistic sensitivity o Sensitivity to formal symbolic notations o Sensitivity to specialized terminology • Meta-linguistic sensitivity o Symbolic distance between sign and object o Sensitivity to the structure and logic of mathematical language	

Table 1 - Sieprinska et al.'s (2002) model for TT

Although the model was developed and used for a study involving Linear Algebra, it was also used in studies concerning different areas of mathematics and often for purposes different than that in Sierpinska et. al's (2002) study; for instance, to describe students' reasoning about real numbers (Zachariades et. al, 2012), to investigate students' TT about absolute value inequalities (Sierpinska et. al, 2011), and even as a basis of TT for developing a model for understanding fractions at the elementary school level (Nicolaou & Pitta-Pantazi, 2010). In the current study the model was used to investigate students' engagement in TT in a Calculus class.

While the model displayed above stands alone as a characterization of TT, an *operationalization* of the model makes it a working 'tool' to identify occurrences of TT in empirical research. In our operationalization, we follow Sierpinska et al. (2002); in their study, the model was operationalized by defining a set of *theoretical behaviors* (TBs) associated to each type of TT. For example, *generalizing a solution* is a TB associated to *reflective thinking*; *modeling a problem* or *referring to previously learned concepts* are TBs associated to systemic thinking.

The display of a TB by a subject is to be understood as an indication of the subject's engagement in the corresponding type of TT. Since TBs that are pertinent to one context may not be to another, the operationalization of the model may differ depending on the context of the study. Therefore, although our operationalization follows closely that of Sierpinska et al. (2002), it has been customized to the context of an introductory Calculus course.

To determine the set of TBs which would operationalize the model for our study and accurately describe the types of behavior we might observe in our data, we referred to the questions we created for the study. For each question we imagined answers that would have involved thinking theoretically; in those answers we identified *features of discourse* that would act as indicators of a certain type of TB. For example, in our operationalization, *writing a general statement* is a feature of discourse indicative of the TB *generalizing a solution*, which in turn is indicative of *reflective thinking*. In Table 2 we show examples of features of discourse associated to the TB *generalizing a solution*.

Category of TT	TB	Features of discourse
Reflective	Generalizing a solution	• Writing a single *general* statement for positive and negative series by considering the absolute value of terms. • Indicating that integral $[a,\infty)$ is convergent for *all* $a\geq 1$. • Indicating that the integral of g is convergent over *any* subinterval of $[1, \infty)$. • Remarking that the addition of *any* non-zero constant to the integrand would result in a diverging integral.

Table 2 - Sample of how our model is operationalized with features of discourse and TBs.

We defined 13 TBs and 68 features of discourse; due to space constraints, we display the operationalized model (Table 3) with only some examples of TBs.

Feature of TT	Samples of corresponding TBs
Reflective	Generalizing a solution
Systemic • Definitional • Proving • Hypothetical • Contextual	 Referring to definitions when deciding upon meaning Engaging in a proving or reasoning activity Being aware of the conditional character of a mathematical statement Modeling a problem
Analytic	Being sensitive to logical connectives, particularly to implication and its negation

Table 3 - Our model of TT, including sample TBs

Methodology

The study was carried out in an *Integral Calculus* course stretched over one term (thirteen weeks) with an average of thirty five students attending the two classes (1h15 each) per week (students in this course have previously passed a *Differential Calculus* course). The instrument used to engage students in TT was a set of questions related to material previously covered in the course, each presented to students in the form of an optional quiz. Each quiz consisted of one or two questions that were designed in a way such that meaningfully answering a question would require students to think theoretically. Once a week, at the end of one of the two sessions, students were given a quiz and fifteen minutes of the class time to respond to the question. The instructor had to manage the class time to allow for the quizzes to take place while keeping up with the weekly objectives stated in the course outline (set up by the course examiner). Although this was a challenge, the instructor managed to keep the schedule; at the end of the semester, the class average was similar to the other sections and in the course evaluation, students praised the quizzes and the opportunities these provided (see *Conclusions* section).

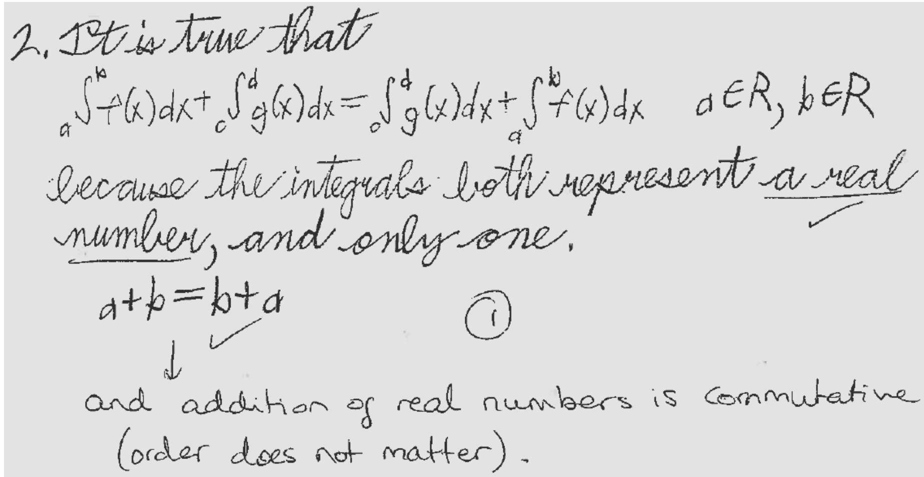
The instructor explained to students that taking the quiz was optional and that they would be awarded with *bonus marks* (up to 5%) on their course grade for a complete response or an incomplete response containing valid arguments. Quiz questions were of a conceptual nature, aimed at prompting a type of behavior that we could characterize as a display of TT. There was usually not a single solution path that had to be followed, but students were asked to "justify" their answers and generally be as expressive as possible. We show three of the quiz questions in the table below.

Q2	Q7	Q11
Is it true that $\int_a^b f(x)dx + \int_c^d g(x)dx = \int_c^d g(x)dx + \int_a^b f(x)dx$ where a, b, c, and d are real numbers?	Explain, in your own words, why this theorem is true: If the series $\sum_{n=1}^{\infty} a_n$ is convergent, then $\lim_{n\to\infty} a_n = 0$	If $g(x)$ is continuous for all real numbers and $\int_1^{\infty} g(x)dx$ is convergent, is $\int_{10}^{\infty} g(x)dx$ also convergent?

Table 4 - Examples of quiz questions (from left to right: Questions 2, 7, and 11)

A response was awarded 1 point if it was valid or 0.5 if it was incomplete but contained elements of sound reasoning. In some cases, even if the response was incorrect but displayed evidence of an awareness of the requirement of the task, 0.5 was awarded as well. Otherwise, if the response was incorrect and did not contain any valid arguments, or if no response was given, no points were awarded.

Quizzes were returned with a grade, and minor corrections and suggestions for improving the quality of their responses when they were inadequate. For instance, if a response contained incorrect reasoning, the instructor provided counter-examples or suggestions as to why the reasoning may not be sound. Students were not provided with complete answers to the questions to allow them the freedom of developing their own style in answering the questions, especially considering that there was usually not a single correct answer to a question. Instead, students were encouraged to review the feedback they received and refer to the instructor to further explore questions and their responses. Copies of students' responses were kept for analysis. Examples of responses to questions (displayed in Table 4 above) that received partial or full credit are shown below to illustrate the grading and type of feedback that was given.



2. It is true that

$\int_a^b f(x)dx + \int_c^d g(x)dx = \int_c^d g(x)dx + \int_a^b f(x)dx$ $a \in R, b \in R$

because the integrals both represent a real number, and only one.

$a+b=b+a$

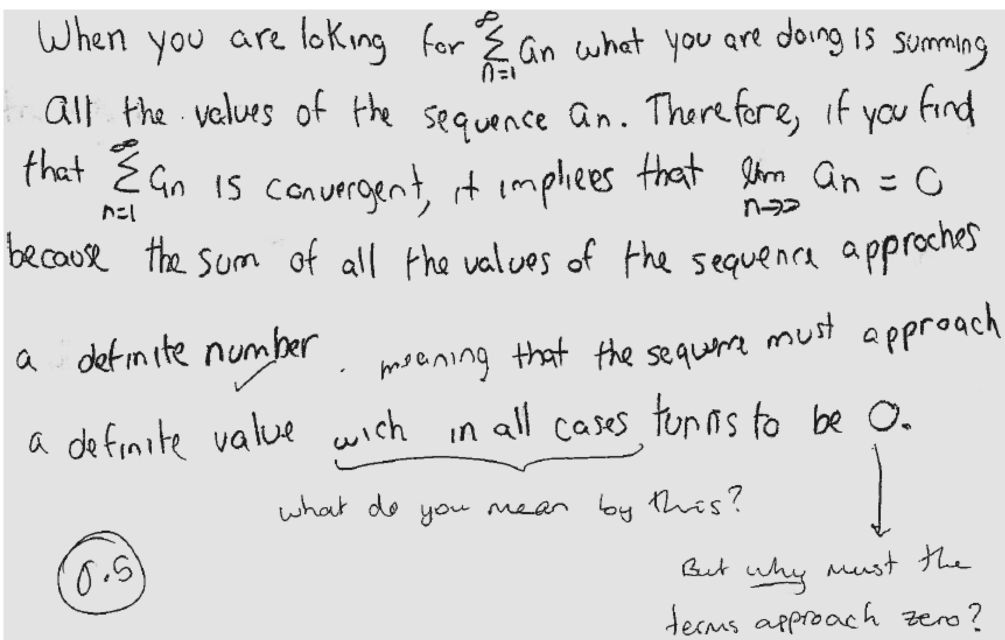
1

and addition of real numbers is commutative (order does not matter).

Figure 1 - Example of a response to Q2 which received full credit and the feedback given

The student whose response is displayed in Figure 1 received full credit but also a suggestion to improve the quality of his answer; in particular, the use of mathematical terminology to describe a property he referred to.

The response in Figure 2 correctly explains the premise of the problem but does not justify why the conclusion must necessarily be true. The response was thus awarded partial credit.



When you are loking for $\sum_{n=1}^{\infty} a_n$ what you are doing is summing all the values of the sequence an. Therefore, if you find that $\sum_{n=1}^{\infty} a_n$ is convergent, it impliees that $\lim_{n\to\infty} a_n = 0$ because the sum of all the values of the sequence approches a definite number. meaning that the sequence must approach a definite value wich in all cases turns to be 0.

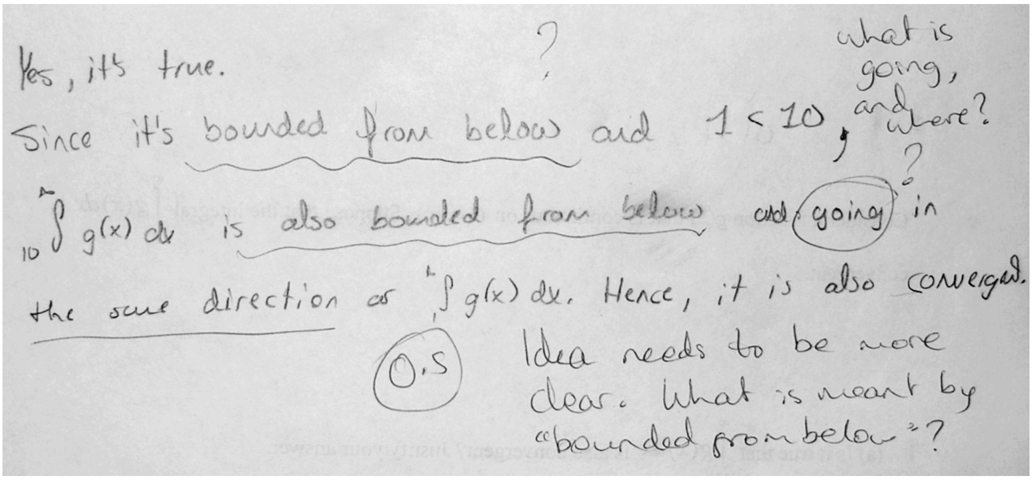
what do you mean by this?

0.5

But why must the terms approach zero?

Figure 2 - Example of a response to Q7 which received partial credit and the feedback given

The response to Q11 in Figure 3 contains elements of sound reasoning; the student states that "$1 < 10$" and is perhaps referring to the (finite) value of each integral. The response is lacking rigor, however, and received partial credit.



Yes, it's true. ?

Since it's bounded from below and $1 < 10$, $\int_{10}^{\infty} g(x)\,dx$ is also bounded from below and going in the same direction as $\int_{1}^{\infty} g(x)\,dx$. Hence, it is also converged.

what is going, and where?

?

0.5 Idea needs to be more clear. What is meant by "bounded from below"?

Figure 3 - Example of a response to Q11 which received partial credit and the feedback given

Three types of analyses were designed: *Question*, *Class*, and *Student analysis*. The *Question analysis* is an a posteriori analysis of the questions; the goal is to revise our prediction of the types of TT that each question invites. This analysis is essential for the design of future iterations. The *Class analysis* would uncover how the class *actually* engaged in TT in answering each question. The *Student analysis* would reveal how every student engaged in TT across the entire set of quizzes. While the latter two analyses are performed in similar ways, they are each valuable in the different results that they offer. Due to space constraints and the depth involved in discussing the results of each type of analysis, the results given in the current paper will be limited to those of the *Class analysis*.

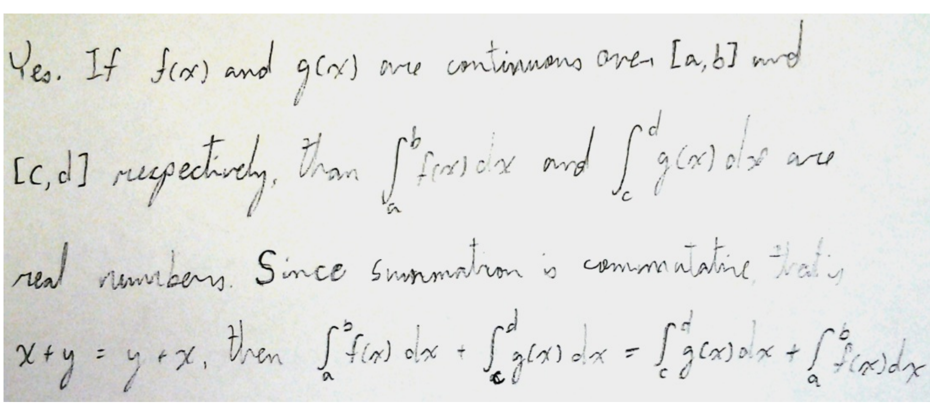
For each question, we searched, across all participants, for displays of TBs (indicated by features of discourse) in students' responses. We marked each display of a TB (per student) as well as the total number of displays of each TB, and feature and category of TT. The results are given in the next section.

Results and Analysis

We begin this section by providing examples of students' responses in which we could identify occurrences of TT; general results of the *Class analysis* follow.

Examples of reflective *thinking:*

Figure 4 displays a student's answer to Q2 (displayed in Table 4 above).



Yes. If $f(x)$ and $g(x)$ are continuous over $[a,b]$ and $[c,d]$ respectively, then $\int_a^b f(x)\,dx$ and $\int_c^d g(x)\,dx$ are real numbers. Since summation is commutative, that is $x+y = y+x$, then $\int_a^b f(x)\,dx + \int_c^d g(x)\,dx = \int_c^d g(x)\,dx + \int_a^b f(x)\,dx$

Figure 4 - Response 1 to Q2

The statement of the problem does not discuss the existence of the integral nor the continuity of the functions over the intervals [a,b] and [c,d] respectively; in fact, in the way that the problem is posed, the emphasis is on the re-ordering of the integrals, and it seems quite clear that the question 'allows' the reader to assume these hypotheses or not even consider them. Yet, this student took the initiative to recall the conditions which would allow him to then proceed with the identification of the integral with a real number; in the context of the question, we consider that the student provided additional detail and displayed what we characterize as "an investigative ("researcher's") attitude towards mathematical problems"- a type of behavior which we associate with *reflective* thinking.

Examples of systemic *thinking:*

Figure 5 displays a student's response to Q2 (see Table 4). We identified two theoretical behaviors in this response: The student first associated the definite Riemann integral with a real number, thus relating two different concepts within a system – a behavior associated to *systemic* thinking.

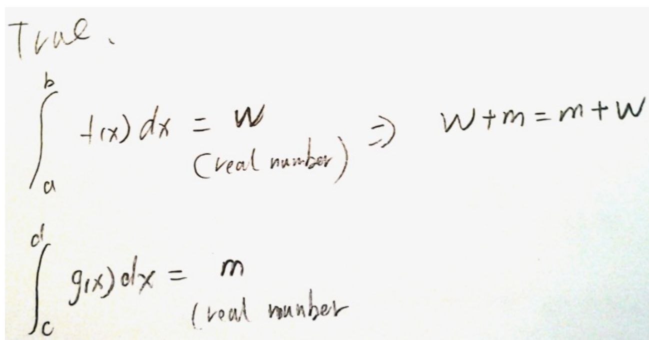
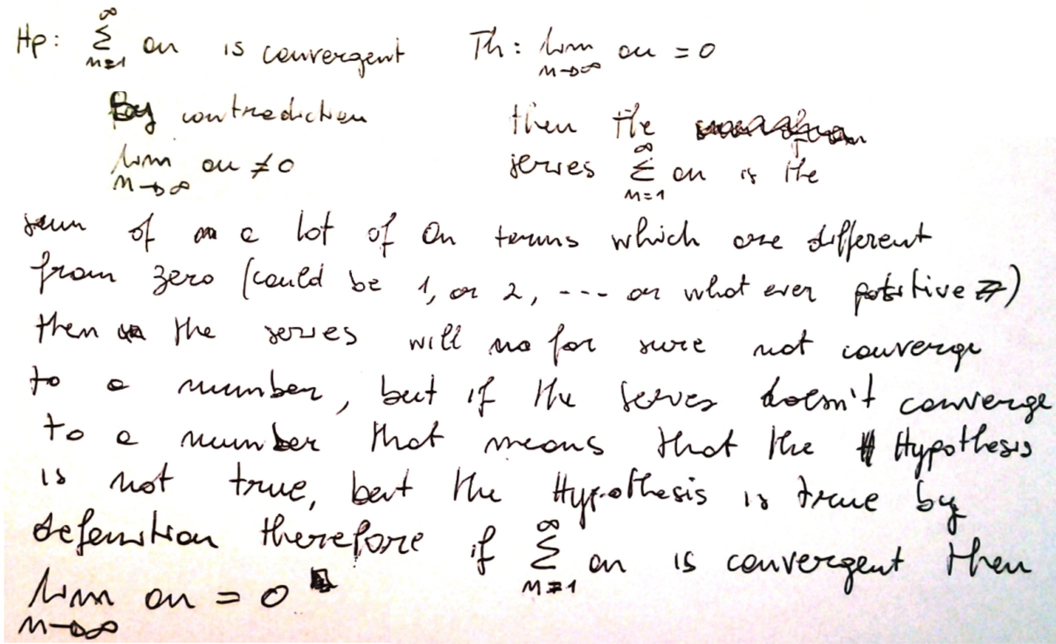


Figure 5 - Response 1 to Q2

The student then used the commutative property of addition of real numbers to establish the sought equality; again a behavior associated with *systemic* thinking as it involves referring to properties of an operation (addition) on objects in a particular system (the set of real numbers).

In his response to Q7 (see Table 4) a student argued about the truth of a statement by reasoning by contradiction (Figure 6):


Hp: $\sum_{n=1}^{\infty} a_n$ is convergent Th: $\lim_{n\to\infty} a_n = 0$

~~By~~ contradiction

$\lim_{n\to\infty} a_n \neq 0$

then the series $\sum_{n=1}^{\infty} a_n$ is the sum of a lot of a_n terms which are different from zero (could be 1, or 2, --- or what ever positive #) then the series will ~~no~~ for sure not converge to a number, but if the series doesn't converge to a number that means that the ~~#~~ Hypothesis is not true, but the Hypothesis is true by definition therefore if $\sum_{n=1}^{\infty} a_n$ is convergent then $\lim_{n\to\infty} a_n = 0$

Figure 6 - Response 1 to Q7

In another response to this question, which we will call *Response 2 to Q7*, a student wrote "If $\sum_{n=1}^{\infty} a_n$ converges to a number, it means that it has to be adding smaller and smaller numbers for it to be able to converge. If we said $\sum_{n=1}^{\infty} a_n$ converges and $\lim_{n\to\infty} a_n \neq 0$, that would not be true. It would be diverging." While Response *1* is written more rigorously, both Responses *1* and *2* contain elements of 'proving by contradiction'. In each of these answers we identified *reasoning* activities- a behavior indicative of *systemic* thinking.

Examples of analytic *thinking:*

Q6 (Figure 7) requires one to explain why the value of the area under the second curve is equivalent to the value of the area under the first. In her response, one student used a 'substitution' with $u = x^2 - 4$ to transform $\int_2^5 x(x^2-4)\mathrm{dx}$ into $\int_0^{21} \frac{u}{2}\mathrm{du}$. She then correctly remarked that this integral is actually equivalent to the integral $\int_0^{21} \frac{x}{2}\mathrm{dx}$ given in the problem since "u is just a variable". This student was able to distance the symbol 'u' from its meaning- a theoretical behavior which is associated with *analytic* thinking.

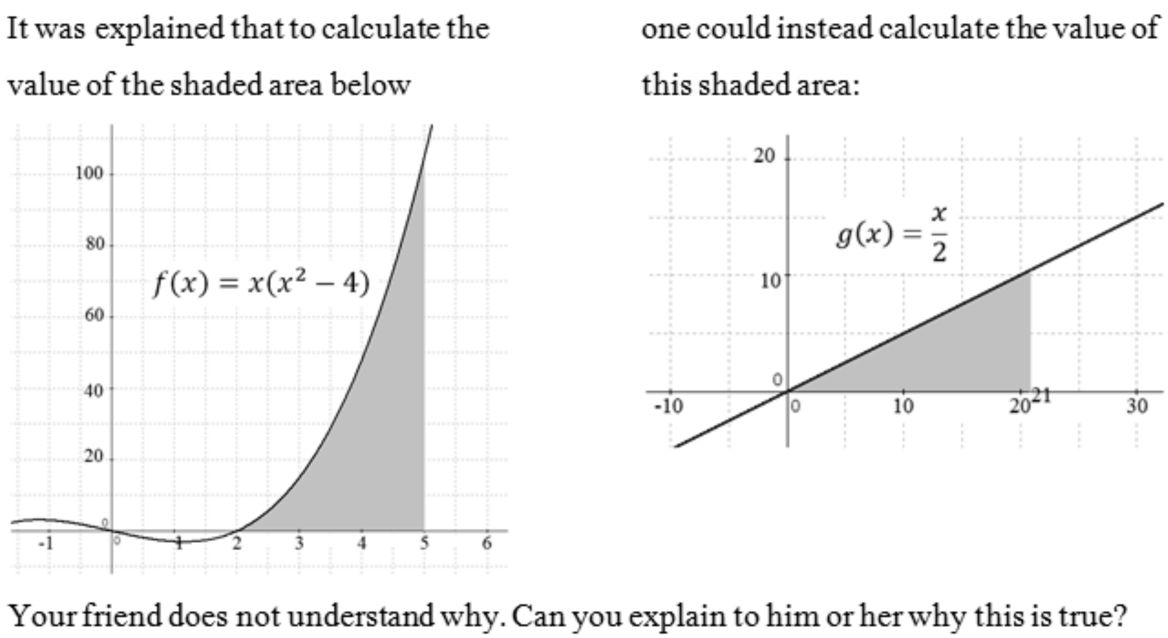


Figure 7 - Q6

Reslults of the Class analysis*:*

We carried out the *Class analysis* by analyzing student responses to each question at a time. We marked the number of times that each TB, and thus each category of TT, was displayed in student responses to each question. This gave us the number of students who chose to participate in the quizzes, the number of students who actually engaged in TT while responding, and the number of occurrences of TT while responding. The number of occurrences of TT was typically higher than the number of students who engaged in TT since one student could engage in more than one type of thinking. With these results we constructed Table 5 which will be the center of our discussion in the next few paragraphs.

The second column of the table displays the number of students who chose to respond to the corresponding question. Fewer and fewer students participated throughout the semester, which is not surprising since typically fewer students attend class as the semester progresses; however, what we found remarkable is that every student who was present on a ‘quiz day’ chose to take the quiz. For us this was a significant result, especially given that the quizzes were carried out at the end of the class and students were aware that they could leave (early) and had no obligation to take the quiz; furthermore, while some students continually received credit for valid responses, some consistently did not – yet they chose to continue taking the quizzes anyway. Moreover, many students were curious to know the correct responses to the questions and often stayed after class to discuss these with the instructor.

Question	# Participated	# (%) Engaged in TT	# Occurrences of TT
Q1	47	33 (70%)	59
Q2	47	28 (60%)	43
Q3	41	24 (59%)	25
Q4	41	34 (83%)	47
Q5	40	21 (42%)	23
Q6	39	27 (69%)	60
Q7	33	23 (70%)	33
Q8	35	12 (34%)	20
Q9	35	12 (34%)	14
Q10	29	23 (59%)	32
Q11	27	22 (81%)	32
Q12	27	10 (37%)	10

Table 5 - Count of participation, engagement in TT, and occurences of TT

The third column of Table 5 displays the number of students who engaged in TT (versus the number who participated, in the second column). The results show that the questions were successful in engaging students in TT; in fact, as shown by the percentages in column three, a significant portion of the class engaged in TT throughout the quizzes

The last column indicates the number of occurrences of TT per question which was typically significantly higher than the number of students who engaged in TT; for instance, while 28 students were engaged in TT in Q2, we identified 43 occurrences of TT – meaning that several students likely engaged in more than one type of thinking, as was the case in the response in Figure 5.

Comparing results shown in the second, third, and fourth columns provides more global information about the questions. For instance, 41 students participated in both Q3 and Q4; however, 34 out of these engaged in TT in Q4, but only 24 in Q3. Furthermore, the number of occurrences of TT is 47 for Q4, compared with only 25 for Q3, likely indicating that several students engaged in more than one type of thinking in Q4 but not in Q3. Similar conclusions can be drawn when comparing the results of Q11 and Q12: Despite the same number of participants, many more students were actually engaged in TT in Q11, and many were likely engaged in more than one type of TT as well. While such results might not be sufficient to compare questions to each other, they inform us that the questions did not equally engage students in TT, and that some questions were perhaps more successful at engaging students in TT than others.

A closer examination of the results of the *Class analysis* revealed that the highest occurrence of TT was *systemic* (see Table 6 below). This result could be quite expected given the context of the study since topics and problems pertinent to Calculus involve activities which most often motivate behaviors similar to those associated with systemic thinking; we would expect that a similar study in a Linear Algebra class, for instance, would reveal a higher occurrence of analytic thinking. In fact, as displayed in Table 6, analytic thinking was seldom invited by the question (N/A indicating that the type of thinking was not invited by a question), and students actually engaged in analytic thinking only in Q6 (4 occurrences). The total number of occurrences of each category of TT per question is what constitutes the last column of Table 5. A question that might naturally be asked at this point is whether it might be of interest to design questions with the intention to invite particular types of thinking, or particular theoretical behaviors.

Question	Reflective	Systemic	Analytic
Q1	19	40	N/A
Q2	14	29	N/A
Q3	6	19	N/A
Q4	2	45	N/A
Q5	12	11	N/A
Q6	26	30	4
Q7	8	25	0
Q8	1	19	N/A
Q9	2	12	N/A
Q10	0	32	N/A
Q11	2	30	N/A
Q12	0	10	N/A
Total	**92**	**302**	**4**

Table 6 - Occurrence of reflective, systemic, and analytic thinking in student responses

Conclusions

College-level Calculus instructors are often compelled to certain institutional constraints; in this study we do not discuss alternative situations in which these constraints are loosened or eliminated- we take on a different perspective. Our study shows that by posing additional non-routine tasks chosen in a way to promote theoretical thinking, creating a space in which students can actively engage in theoretical thinking (should they wish to do so) is indeed possible, despite these constraints; thus incorporating what we believe to be an essential part of a Calculus course: a theoretical thinking component.

While designing this study the researchers and instructor made an ethical decision: they agreed that taking 15 minutes away of class every week would be worthwhile to run the quizzes. As pointed out above, the instructor followed the schedule stated in the course outline and the class average was similar to the other sections (among the best 3 out of 6). Furthermore, in the course evaluations (run by the university), students praised the quizzes: "The quizzes provide excellent feedback on our understanding of the theory in class" and "I love the quizzes because they test your knowledge without consequences". For us this, together with the results showed in the previous section, constitutes a strong indication of the success of the tool to engage students in TT. Of course, the results call for further research; some questions that this work might seek to answer are proposed below.

In the study some questions proved to engage students in TT more than others; what features of a question make it more TT-engaging? We conjecture that these features are not only mathematical (i.e., related to the mathematics required to answer the question) but didactic (concern the level of difficulty of the question in relation to the mathematical maturity of the students). Students in this study engaged in some categories and features of TT more than in others; will Calculus students typically engage more in those categories and features? If so, why would this be the case? A closer look at this question might contribute to our understanding of the TT involved in the learning of different mathematics concepts. When refining and modifying the questions for an iteration of the study, one can place emphasis on engaging students in a *variety* of categories and features of TT or on some *particular* ones. These approaches might give different results, and shed light on different aspects of TT and students engagement in it – which we believe are worth exploring.

References

Barbé, J., Bosch, M., Espinoza, L., & Gascón, J. (2005). Didactic restrictions on the teacher's practice: The case of limits of functions in Spanish high schools. *Beyond the apparent banality of the mathematics classroom*, 235-268.

Boesen, J., Lithner, J., & Palm, T. (2010). The relation between types of assessment tasks and the mathematical reasoning students use. *Educational studies in mathematics*, *75*(1), 89-105.

Hardy, N. (2009). Students' perceptions of institutional practices: The case of limits of functions in college level Calculus courses. *Educational Studies in Mathematics, 72*(3): 341-358.

Hardy, N. (2009). Institutional practices and conceptual development - a case study of college level calculus courses. *Didactica Mathematicae, 32*: 93-119.

Hardy, N., & Challita, D. (2012). Students' perceptions of the role of theory and examples in college level mathematics. Accepted. *Proceedings of the Psychology in Mathematics Education North American Chapter*, Kalamazoo, Michigan, November 1-4, 2012.

Hardy, N., & Sierpinska, A. (2011). Mathematical organization of some French and English algebra textbooks used in remedial courses in colleges and universities in North America. *Proceedings of the Colloquium "Formation à la recherche en didactique des maths"*, UQAM, 24-26 March, 2011, pp. 239-250.

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

Nicolaou, A. A., & Pitta-Pantazi, D. (2010). A new theoretical model for understanding fractions at the elementary school. *University of Cyprus*.

Raman, M. (2004). Epistemological messages conveyed by three high-school and college mathematics textbooks. *The Journal of Mathematical Behavior, 23*(4), 389-404.

Schoenfeld, A. H. (1987). What's All the Fuss About Metacognition?. *Cognitive science and mathematics education*, 189.

Selden, A., Selden, J., Hauk, S., & Mason, A. (1999). *Do Calculus students eventually learn to solve non-routine problems*. Technical report, Department of Mathematics, Tennessee Technological University.

Sierpinska, A., Nnadozie, A., & Oktac, A. (2002). A study of relationships between theoretical thinking and high achievement in Linear Algebra. Unpublished manuscript. Retreieved from http://www.annasierpinska.wkrib.com/index.php?page=publications

Sierpinska, A., Bobos, G., & Pruncut, A. (2011). Teaching absolute value inequalities to mature students. *Educational Studies in Mathematics*, *78*(3), 275-305.

Vygotsky, L. S. (1987). LS Vygotsky, collected works: Volume 1, Problems of general psychology, including the volume *Thinking and Speech*, Plenum Press, New York and London.

Zachariades, T., Christou, C., & Pitta-Pantazi, D. (2012). Reflective, systemic and analytic thinking in real numbers. *Educational Studies in Mathematics*, 1-18.

STRUGGLING WITH CONFIDENCE: PERSISTENCE IN MATHEMATICS AMONG SECONDARY MATHEMATICS MAJORS

Joe Champion
Texas A&M University - Corpus Christi

Ann Wheeler
Texas Woman's University

Completing undergraduate mathematics courses is a central feature of the professional preparation of most prospective secondary mathematics teachers. However, there is little mixed methods research into the patterns of course taking, performance, and persistence among mathematics majors, in general, and among secondary mathematics majors, in particular. Drawing from a sample of 42,825 mathematics enrollment records at two universities over a six-year period, this study uses a social cognitive perspective to better understand mathematics majors' performance and persistence in undergraduate mathematics courses. Quantitative analysis is accompanied by case studies gleaned from exploratory qualitative interviews of nine secondary mathematics majors at one of the universities. Implications include potential strategies for mathematics programs and faculty to support the success of secondary mathematics majors in undergraduate mathematics coursework.

Key words: undergraduate mathematics performance, academic persistence, secondary mathematics teacher preparation

In a recent report, an advisory panel to the U.S. president called for a $100 million national experiment addressing the mathematics preparation of Science, Technology, Engineering, and Mathematics (STEM) students (PCAST, 2012). The recommendations include expanding mathematics instruction by faculty from other disciplines (e.g., physics or engineering) and building "a new pipeline for producing K-12 mathematics teachers from undergraduate and graduate programs in mathematics-intensive fields other than mathematics." (PCAST, p. vii). While questioning the prominent role of mathematicians in the preparation of STEM professionals teachers (AMS, 2012), the report further suggests students in a range of STEM programs may be better served by learning the specialized mathematics needed for their fields from faculty in those fields. The PCAST recommendations challenge the long-standing teacher education practice in which mathematicians play a primary role in the preparation of secondary mathematics teachers (CBMS, 2012).

One way to address the PCAST recommendations for preparing secondary mathematics teachers is to consider the performance of mathematics majors in mathematics courses. However, there is relatively little research addressing questions such as: How well do mathematics majors perform in undergraduate mathematics courses? What proportion of mathematics majors continue in mathematics despite one or more failing course grades to complete their degree requirements? What sources of support do mathematics majors utilize if/when they struggle in mathematics courses? This study provides insights into these questions through a rare mixed-methods investigation on the performance and persistence of mathematics majors in undergraduate mathematics courses. The design focuses on quantitative analysis of course enrollments and letter grades in mathematics courses among a large sample of mathematics majors. In addition, we present exploratory case studies designed to contextualize the numerical data by describing the lived nature of mathematics persistence among a small subsample of secondary mathematics majors.

The Mathematical Preparation of Secondary Mathematics Teachers

A U.S. undergraduate who plans to become a secondary (grades 6-12) mathematics teacher must complete a federal requirement that all secondary teachers earn “highly qualified” status (NCLB, 2001) through state certification requirements, earning a bachelor's degree, and demonstrating *subject matter expertise*. In mathematics, subject matter expertise is typically interpreted as including both mathematical content knowledge (MCK) and pedagogical content knowledge (PCK) (Hill, Rowland, & Ball, 2005). While teacher education programs have diverse approaches to building prospective teachers' PCK (e.g., supervised student teaching), nearly all programs address MCK through courses in an undergraduate mathematics major, such as single- and multi-variate calculus, differential equations, linear algebra, introduction to proof, statistics, and abstract algebra (Conference Board of Mathematical Sciences, 2012) . Though it is unclear whether these undergraduate mathematics courses necessarily improve the specialized knowledge of mathematics needed in middle and high school classrooms (Hill et al., 2005; Kahan et al., 2003), U.S. secondary mathematics teachers typically complete at least four times as many courses in mathematics as they do in education (Monk, 1994).

Mathematics majors collectively consist of just 1% (12,363 of 1,119,579) of bachelor’s degree earners in the U.S. (Lutzer, 2002), and there has been limited research on the patterns of academic outcomes among the even smaller number of students concentrating in secondary mathematics teaching. Existing research does, however, suggest earning a bachelor’s degree in mathematics can be a difficult experience. Mathematics courses have the second lowest grade distribution among all majors (Rask, 2010), and more than 3 of 5 students who declare a mathematics major switch out of the major (Seymour & Hewitt, 1997). Moreover, students' early work toward earning mathematics degrees may be complicated by the role of mathematics courses as a filter for business and STEM programs (Moore & Shulock, 2009) and a variety of social factors, such as a lack of family support and the beliefs about the nature of mathematics (Ma & Kishor, 1997). Secondary mathematics majors may also face the challenge of navigating between grading norms in mathematics courses and those established in education courses, which have the highest grade distribution among all college majors (Rask, 2010). Even a small number of low letter grades in required courses can greatly decrease the likelihood of completing an undergraduate degree (DesJardins, McCall, Ahlburg, & Moye, 2002), where a prospective secondary teacher may perceive poor grades in mathematics courses as foreshadowing future difficulties in completing his or her degree.

Increasing the Secondary Mathematics Workforce

Mathematics is just one of several technical fields (e.g., physics) for which earning an undergraduate degree could be described as both difficult and rare. However, a low supply of undergraduate mathematics degree earners may limit the availability of new secondary mathematics teachers and could potentially limit over 30 years of efforts to address a persistent shortage of secondary mathematics teachers (Ingersoll, 2003; U.S. Department of Education, 2009). The PCAST (2012) recommendations and the $250 million public-private partnership *Educate to Innovate* to add at least 10,000 mathematics and science teachers are just two recent efforts to boost a current U.S. work force of about 128,500 secondary mathematics teachers (Morton, Peltola, Hurwitz, Orlofsky, Strizek, & Gruber, 2008).

Though learning mathematics is central to learning to teach mathematics (Cooney & Wiegel, 2003), there has been limited research into the academic performance of prospective secondary mathematics teachers in undergraduate mathematics courses. This may be due in part to the relatively small number of secondary mathematics majors one can expect to find in an undergraduate mathematics program. Students majoring in secondary mathematics are typically outnumbered in nearly all required courses by STEM majors outside of mathematics (Lutzer, Rodi, Kirkman, & Maxwell, 2007). In 2005, there were just 3,400 secondary mathematics majors and 12,000 applied or liberal arts mathematics majors among over 699,000 students enrolled in undergraduate mathematics courses at the calculus level or beyond (Lutzer et al.). While there have been (mostly qualitative) investigations into the subject matter knowledge gained by secondary mathematics majors through undergraduate mathematics programs (e.g., Bryan, 1999; Even, 1993), there is a need for research into the academic experiences of prospective mathematics majors in their mathematics coursework.

Conceptual Framework

Though the research is couched in the importance of understanding mathematics majors performance from the perspective of teacher preparation, the conceptual perspective is based on a social cognitive (Bandura, 1997) view that *self-efficacy* (one's cognitive assessment of his or her capacity to affect specific outcomes under specific circumstances) and *calibration* (realistic feelings of confidence) have important mediating effects on the relationship between past performance and future performance in academic settings. In the specific domain of mathematics, statistical modeling has pointed to self-efficacy and calibration as having independent and approximately equal effects on mathematics performance (Champion, 2010; Chen, 2002; Hackett & Betz, 1986). However, more social cognitive research is needed on experiences of *sources of self-efficacy* (Usher & Pajares, 2008) in undergraduate mathematics as well as the relationship between mathematics performance and *persistence*, which is marked by continued or increased effort following self-perceived struggles (Zeldin & Pajares, 2000).

Research Questions

The basic goal of this research was to investigate (a) the patterns of course enrollment and letter grades earned by mathematics majors in mathematics courses, and (b) how secondary mathematics majors experience and respond to difficulties in undergraduate mathematics courses. This article addresses these issues by combining quantitative and qualitative data sources to investigate two research questions:

1. (Quantitative) What characterizes mathematics majors' academic performance and persistence in undergraduate mathematics courses at two universities?
2. (Qualitative) What describes prospective secondary mathematics teachers' academic struggles and sources of support in undergraduate mathematics courses at a university?

Methods

Setting

The two universities included in the quantitative strand are similar in many ways – each is a minority serving public institution located in the same central U.S. state, enrolls about 7,500 full-time undergraduate students, admits about 90% of applicants (many of which are community-college transfers), and offers undergraduate and master's programs in mathematics. The main contrasts are that University A is a regional campus in a mid-sized city, while University B is in a large metropolitan area and enrolls mostly female students.

Table 1 summarizes the mathematics program requirements at the two sites. The universities both offer core mathematics requirements completed by both teaching and non-teaching concentrations in their mathematics programs. However, the core requirements at the two sites differ in important ways: (a) University A's mathematics major includes more required upper level mathematics courses and proof-based classes than University B, and (b) University B has electives in their core requirements allowing majors to complete upper division mathematics courses with a mathematics education focus (e.g., Geometry in the Classroom).

Table 1. Core requirements of degree requirements at the two research sites.

	University A	University B
Similar (21 credits)	Calculus 1, 2, 3 (Fr, So, Jr)	Calculus 1, 2, 3 (Fr, So, Jr)
	Linear Algebra (Jr)	Linear Algebra or Matrix Methods (Jr)
	Abstract Algebra (Jr)	Abstract Algebra (Jr)
	Probability & Statistics (Jr)	Statistics (either Fr or Sr level)
Different (16 credits)	Discrete Math I (So)	Math for Liberal Arts (Fr)
	Differential Equations (Jr)	Trigonometry (Fr)
	Introduction to Proof (Jr)	Math Elective (any level)
	Applied Modeling (Sr)	Math Elective (any level)
	Introduction to Analysis (Sr)	Math Elective (Jr or Sr)

Note: course levels: "Fr" = Freshman , "So" = Sophomore, "Jr" = Junior, "Sr" = Senior

Data collection

Data collection consisted of institutional academic records from both universities as well as face-to-face interview and survey data from University B. Quantitative data included student, course, and performance variables for all students enrolled in one or more mathematics courses during the study period of six academic years (fall 2005 to spring 2011). For each enrolled mathematics student, the institutional records included unique student identifier, university, ethnicity, sex, ACT/SAT mathematics score, high school grade point average, age, undergraduate major, academic level, course name, course section, instructor, term, and final letter grade.

The qualitative strand of the study employed semi-structured interviews crafted after the protocol of Seymour and Hewitt's (1997, p. 401) large-scale study of switching- and non-switching undergraduate science, mathematics, and engineering majors. Interview questions invited study participants to discuss experiences surrounding choosing a college major, performance in high school and college mathematics, exam and course grades in mathematics, quality of mathematics instruction, sources of academic support, (self- and peer) experiences of major switching, and career plans. The protocol combined the framework of Bandura's (1997) social cognitive theory of learning with Seymour & Hewitt's themes; this afforded opportunities to address academic performance in the context of self-efficacy and persistence constructs while utilizing the personal and social themes undergraduates typically use to describe their experiences in undergraduate mathematics courses. As part of the interview protocol, participants completed a modified version of the 30-item Mathematics Self-Efficacy and Anxiety Questionnaire (May, 2009) and self-reported their mathematics self-efficacy in the content of each of their undergraduate mathematics courses on a scale of 0 (not confident) to 10 (very confident). Participants also completed a brief background survey for high school mathematics performance and self-efficacy among students who have completed advanced undergraduate mathematics courses (Champion, 2010).

Sample

The quantitative data sample initially included all $N = 42, 825$ enrollments in mathematics courses by undergraduate students at the two universities during the six-year study period. While this larger sample allowed for analysis of course enrollment patterns, measures of persistence in mathematics were focused on a subsample of $n = 12{,}522$ mathematics course enrollments by students who completed all their mathematics classes during the six-year study period. Of this persistence sample, 55% of enrollments were from University A and 45% were from University B. Demographics in the overall sample suggests the data set included more female (72%) than male (28%) students, and no majority student ethnicity (48% White, 31% Hispanic, 14% Black, 4% Asian, 2% Other). In addition, the sample included comparable enrollments by Freshmen (27%), Sophomores (17%), Juniors (20%), and Seniors (26%). Most students were either between the ages of 17 to 22 (48%) or 23 to 27 (35%), with just 16% older than 27.

In order to avoid methodological hazards of "backyard research," the nine qualitative interview participants were purposefully sampled at University B. All undergraduates from University B in the persistence subsample who had previously declared a secondary mathematics major (approximately 40 students) were invited to participate in the study by email, so the interview participants are probably best considered as a self-selected group of secondary mathematics teachers interested in sharing their experiences in mathematics courses. As indicated in Table 2, eight of the participants were still mathematics majors at the time of the interview, and six participants planned on working as a teacher after completing their degree. Just three interview participants planned to teach at the middle- or high-school level.

Table 2. Qualitative summary of interview participants at the time of the interview.

Pseudonym	# Math Courses	Career Plans	Major	Notes
Charlie	10	Actuary	Math	CC transfer, mother
Drew	13	Teach at a CC	Math	mother, learning disability
Ellen	27	Teach Secondary	Math	Mother
Goldie	19	Teach at a CC	Math	CC transfer, mother
Jennifer	18	Undecided	Math	CC transfer, mother
Katherine	4	Teach Elementary	History	9 different majors
Kirsten	10	Undecided	Math	Attended 2 prior universities
Sandra	12	Teach Secondary	Math	CC transfer
Beyonce	12	Teach Secondary	Math	CC transfer

Note: CC = Community College; # Math Courses = number of enrollments in math courses.

Data analysis

In the qualitative strand, the lead author interviewed all participants and administered the surveys to each participant. A student aide created spreadsheets for survey responses and transcribed all the interview data, while the other researcher used the qualitative software package *NVivo* to code the interviews using an open coding (Corbin & Strauss, 2008) strategy to identify emergent themes and characteristics of the variation among participants' experiences in undergraduate mathematics programs. The quantitative analysis of institutional records included standard descriptive and inferential statistical analyses, including cross-tabulations, measures of central tendency, spread, and shape, and tests for marginal differences among performance and persistence measures by student-level and course-level variables.

Results

We first present the quantitative findings in order to build a broad portrait of mathematics performance and persistence among mathematics majors at the two participating universities. The qualitative findings then narrow the focus to the experiences of an information-rich subsample of secondary mathematics majors.

Performance of Mathematics Majors in Mathematics Courses

Mathematics majors represented 1.9% of the unique students who took one or more mathematics courses during the study period. However, enrollments by mathematics majors accounted for 5.8% of all the mathematics enrollments, including at least one mathematics major completed every one of the 28 mathematics course titles offered during the study period. As indicated in Table 3, mathematics majors were found in the highest percentages in classes with the lowest enrollments. For example, mathematics majors formed just 0.8% of the enrollments the 8 introductory general education courses (e.g., College Algebra) that account for 77% (33,177 of 42,825) of the combined mathematics enrollments. Mathematics majors formed less than 10% of enrollments in Trigonometry, Precalculus, and Calculus I, and less than a third of enrollments in Probability & Statistics, Calculus II, and Discrete Math I & II. Mathematics majors were the minority in 6 of the 10 most taken courses by mathematics majors, including Probability & Statistics, Calculus I & II, Discrete Math I, Linear Algebra, and Introductory Statistics I. The course with the highest absolute numbers of enrollments by mathematics majors was Probability & Statistics (224).

Table 3. Majors among students enrolled in undergraduate mathematics courses.

Course Title	Univ.	Student Major at Time of Enrollment							*Total*
		%Busin.	%Educ.	%H. Sci	%LA	%Math	%S/E	Undec.	
Partial Diff. Equations	A	0	0	0	0	100	0	0	*7*
Statistical Theory	A	0	0	0	0	100	0	0	*6*
Mathematics Capstone	A	0	0	0	0	98.3	1.7	0	*60*
Real Analysis	A	1	2	0	1	93.9	2	0	*98*
Introduction to Proof	A/B	2.8	2.8	0.7	2.8	83.8	7	0	*142*
Number Theory	B	0	0	0	0	83.3	16.7	0	*12*
Applied Diff. Equations	A	0	0	0	0	80	20	0	*5*
Abstract Algebra	A/B	1.8	5.6	2.5	4.6	76.8	7.7	1.1	*285*
Differential Equations	A/B	0.8	0	0.8	0.8	66.1	30.5	0.8	*118*
Matrix Methods	B	2.9	4.4	5.9	3.7	64.7	16.9	1.5	*136*
Calculus III	A/B	2.8	3.4	1.6	3.1	55.6	33.2	0.3	*322*
Geometry	A	0	46.1	0	2.3	47.7	3.9	0	*128*
Linear Algebra	A/B	1.5	26.4	0.9	4	41.9	24.9	0.3	*329*
Discrete Math II	A	1.9	0	0.9	0	29.2	67.9	0	*106*
Discrete Math I	A/B	2.3	14.7	1.6	3.4	26.3	51.4	0.3	*617*
Calculus II	A/B	2.7	3.7	5.6	6.1	18.7	62.8	0.5	*863*
Probability & Statistics	A/B	1.3	3.4	22.2	3.6	18.3	51.1	0.1	*1226*

Calculus I	A/B	4.4	5.9	17.3	7.6	7.9	56.3	0.6	*2191*
Precalculus	A/B	4.2	5.7	18.9	8.6	6.5	56	0	*1208*
Trigonometry	A/B	2.5	8.8	34	8.7	5.3	37.3	3.4	*1730*
General Education[a]	A/B	19	13.5	31.6	25.3	0.8	8.7	1	*33177*
Combined		15.4	12.1	28.2	21	5.3	17.1	1	*42825*

Notes: [a] = Basic Algebra, Intermediate Algebra, Math for Liberal Arts, College Algebra, Business Math, Business Calculus, Introductory Statistics I & II; Univ. = University; Busin. = Business, Educ. = Education, H. Sci. = Health Sciences, LA = Liberal Arts, S/E = Science or Engineering, Undec.= Undeclared.

Though mathematics majors were among the least common major types enrolled in mathematics courses (only "Undeclared" had lower absolute enrollment numbers), mathematics majors earned the highest overall distribution of letter grades. Forty-eight percent of mathematics majors maintained an average letter grade of B or better in their mathematics courses with a mean GPA of 2.5 (*Mdn* = 2.9, *SD* = 1.3, Range = 0 to 4). As indicated in Table 4, more than one-in-three enrolled mathematics majors earned a letter grade of A (highest among all major types), and the overall DFWI rate for mathematics majors was 20.8% (lowest among all major types). However, six course titles showed DFWI rates among mathematics majors of the courses of more than 25%, including Abstract Algebra (32%), Introduction to Proof (32%), Linear Algebra (30%), Probability & Statistics (29%), Real Analysis (28%), and Calculus I (26%).

Table 4. Grade distribution in undergraduate mathematics courses by major.

Major Type	%A	%B	%C	%D	%F	%W	%I	Total
Health Sciences	32.6	24.3	18	6.6	11.5	5	2	*11,780*
Liberal Arts	20	23.8	21.4	7.5	14.5	8.7	4.1	*8,843*
Science/Engineering	23	23	22.1	9.9	12.6	8.1	1.3	*7,248*
Business	25.6	24.9	21.5	7.4	10.4	8	2.2	*6,531*
Education	22.7	24.1	21.6	7.4	12.6	6.9	4.6	*5,102*
Mathematics	34.9	26.4	17.9	6.9	8.5	4.7	0.7	*2,249*
Undecided	22.8	29.3	22.8	10.1	11.7	3.4	0	*386*
Combined	26.1	24.2	20.4	7.6	12.1	7	2.6	*42,139*

Persistence among Mathematics Majors in Mathematics Courses

The persistence sample of mathematics course enrollments by students who completed all their mathematics courses during the study period (n = 12,522) included just 98 unique students (0.8%) who declared a mathematics major during the study period. The mean number of mathematics courses attempted by mathematics majors was 5.0 (*Mdn* = 3, SD = 4.5, Range = 1 to 19), and 69.4% of mathematics majors took 5 or fewer mathematics courses. However, just 29.4% of mathematics majors passed 6 or more mathematics classes with a letter grade of C or better (M = 4.3, *Mdn* = 3, SD = 4.4, Range = 1 to 19). More than half (54.1%) of mathematics majors attempted 3 or fewer mathematics courses, 15.3% attempted 4 or 5 mathematics classes, 12.3% attempted between 6 and 9 mathematics classes, and 18.4% attempted 10 or more mathematics classes. Based on the enrollment numbers, we estimate only about one in four mathematics majors could have completed enough mathematics courses to fulfill undergraduate degree requirements.

Table 5 summarizes the distribution of failing letter grades (DWFI) earned by mathematics majors at the research sites. Overall, 35.7% of mathematics majors failed at least 1 mathematics class, including 39.3% of mathematics majors who enrolled in 7 or more mathematics classes. There were 8 courses for which 10% or more of the enrolled mathematics majors had previously attempted the course, including Introduction to Proof (19.3%), Abstract Algebra (16.0%), Real Analysis (15.2%), Probability & Statistics (14.7%), Calculus I (12.2%), and Linear Algebra (10.1%). Overall, 9.5% of mathematics majors enrolled in mathematics courses had previously attempted the course, and 1.3% had previously attempted the course two, three, or four times.

Table 5. Joint distribution of mathematics letter grades among mathematics majors.

	Number of Enrollments in Mathematics Courses							
# D/W/F/I	1	2	3	4	5	6	7+	Combined
% 0	66.7	90	56.3	55.6	50	100	60.7	64.3
% 1	33.3	0	25	22.2	0	0	10.7	18.4
% 2+	0	10	18.8	22.2	50	0	28.6	17.3
Total	*27*	*10*	*16*	*9*	*6*	*2*	*28*	*98*

Struggles in Mathematics Courses among Secondary Mathematics Majors

To help contextualize the performance and persistence data, and to provide some insights into the variety of secondary mathematics majors' experiences in mathematics, the qualitative analysis focused on understanding the mathematical struggles described by interview participants and the diverse ways in which some of the students worked to overcome those obstacles. All of the participants considered earning a letter grade of C or worse as indicative of "struggling in a mathematics course." In all, we found eight areas identified by multiple students as sources of struggles in mathematics courses, including qualities of the course instructor (5 participants), family/parenting issues (4 participants), difficulty of course content (4 participants), inadequate understanding of prerequisite content (3 participants), online course format (3 participants), difficulty in understanding the textbook (2 participants), and a perceived lack of real-world applications (2 participants).

All of the study participants described struggling in a substantial way in at least one of their undergraduate mathematics courses. Five participants failed (earned less than a C) at least one required mathematics course on their first attempt, and three of the participants took a specific mathematics course three or more times. For all but one of the participants (Goldie was the exception), students described first encountering substantial difficulties in a mathematics courses while in college. On the whole, the participants described performing well as elementary and secondary school mathematics students and recalled rarely struggling to learn new mathematical ideas prior to college.

Some students seemed mostly unshaken by failing to pass a mathematics course on the first attempt. Goldie had struggled to earn passing grades in mathematics since childhood but pressed on through remedial mathematics at a community college and several attempts at Calculus I before finding a recipe for earning passing grades through tutoring, group work, and extensive study time. Sandra, who had earned an A in every mathematics class up through Calculus I, classified her frustrations in earning a C in Calculus II as reasonable given difficult content:

> Calc II was difficult. I had a different professor than I had had for Calc I, and I actually enjoyed the way he taught class more… I just felt like he really pushed you to learn more. I ended up with a C in that class just because I felt like there is a lot to learn…. I realized that I hadn't learned as much as I thought I had in Calc I…. I knew

> how interested in Math I was, and how I enjoyed doing it and those kinds of things, so I never really thought just because I made a C I shouldn't pursue what I want to do.

Another study participant, Kirsten, described a great deal of struggle in a series of mathematics courses. Kirsten excelled in high school mathematics courses and initially entered college planning to become a high school mathematics teacher. She earned high grades in calculus and other lower-division college mathematics classes, and, like Katherine, ran into a difficulty in a single course. Kirsten described how earning a D in an introductory Real Analysis course at another university impacted her undergraduate experience:
I had a complete lack of understanding [of Real Analysis]…. I had to retake it because it was a major class, but I just I couldn't go back to school. I didn't know what I wanted to do with my life. That D really stung me a lot. So…. for the fall, I substitute taught, and that was important in getting me back to school, because I did a lot of substitute teaching and I liked the school.

Kirsten's poor performance in Real Analysis led her to temporarily withdraw from the university and return home. She returned to school a semester later and devoted a significant amount of time and effort into Abstract Algebra:

> I worked myself to death …. This is what I call my crazy time. I got so depressed after that class that I actually started seeing a psychologist…. I took Analysis again. I took Abstract Algebra again. I took a Geometry class. I mean these are classes that I took that I pretty much just completely failed. I was trying because I was close to being done…. I kept on trying, thinking I would maybe just be done, but I was emotionally shot at this point. Abstract Algebra killed that.

After failing Abstract Algebra, Real Analysis, and Geometry, Kirsten formally withdrew from the university and took two years to work and seek treatment for clinical depression. She attributed her mathematics hardships as major sources of her depression, recalling, "I'm struggling with math classes. How can I even graduate? How can I become a math teacher? Is this the right path? What am I supposed to do with my life? What do I do now?" Before returning to school, Kirsten described having to work through a fear of failure in mathematics with a therapist. She eventually came to believe, "I probably am capable of graduating college. I am capable of this… Just because I failed once or twice doesn't mean that I will continue to fail in every other thing." Three years after leaving the university, Kristen moved back to her home state and enrolled part-time in one mathematics course at University B. The next semester she took two courses (including Abstract Algebra), and after earning an A in all her classes, described plans to return as a full-time student in the fall. She was unsure of whether or not she still wanted to teach high school mathematics but confidently expected to complete her undergraduate mathematics degree in the near future.

A third participant relayed her experience as a secondary mathematics major taking mathematics courses primary through her experience of earning a C in Calculus II. Katherine took great pride in her academic success in high school, where she made high A's in all of her classes. She entered college intending to become a secondary mathematics teacher, and earned A's in her first three college mathematics courses. After one attempt at Calculus II, though, she decided to switch out of the mathematics major. At the time of the interview, she had switched majors 9 times in the subsequent year and was currently a History major with plans to teach elementary school. Katherine became visibly upset when discussing how Calculus II led her to switch out of the mathematics major:

> I hated that professor. Like, hated. I am serious. He made me feel so stupid, and that is why I changed my major. He made me feel like an idiot. He would work problems,

> and the problems that he worked were not the ones that were on the test. He would explain something, you would understand it for about 5 seconds, and then you would be, like, I still don't understand…. I wish I had never taken that class. It brought down my GPA. That class ruined my GPA. He was, and still is, the worst teacher that I have ever had. He taught more of the core classes that were required for the major, and I didn't want to take them. That's about it.

The preceding stories exemplify just some of the intense personal and academic struggles the secondary mathematics majors recounted as they pursued undergraduate mathematics degrees. Some participants described feeling nearly hopeless while struggling to learn calculus content with a lack of understanding of pre-requisite content. Four participants struggled to catch-up in classes after missing substantial content due to parenting and family issues, including caring for their children, grieving the death of a close relative, and even having a child mid-way through an Abstract Algebra course.

Sources of Support in Mathematics Courses among Secondary Mathematics Majors

Nearly all of the study participants described finding ways to cope with their struggles in mathematics courses, leverage support structures, and persist in mathematics. Table 6 summarizes participants' sources of support as they encountered academic difficulties in undergraduate mathematics courses.

Table 6. Participants' sources of support in mathematics courses.

Source of Support	Counts (n = 9)
Teacher of the Course	9
Tutor (provided or hired)	5
Small Study/Homework Group	4
The Mathematics Lab	4
Family Member or Friend	3
Disability Support Services	1

All participants reported seeking help from their professors, but students' perceptions of those interactions varied. Katherine, whose experience in Calculus II led to her switching out of the secondary mathematics major, did not find her visit to the instructor's office helpful:

> He told me that I needed to study longer and work out more problems…I told him that I was doing the homework, and no matter how hard I tried on the test, I wasn't making the grades that I wanted to make. He told me that I just needed to study more.

Through several students described unhelpful (or even hurtful) comments from instructors when they sought help in mathematics courses, some participants described receiving a significant amount of help and guidance from instructors. For instance, Beyonce attributed much of her success in Abstract Algebra to a teacher who generously helped students:
I feel like [the Abstract Algebra professor] really tried to help you understand. If you didn't get it, she would think of a different way to explain it, and I always went to her office hours to ask her questions. She was always very willing to help.

Four of the nine participants indicated their preferred mode of studying and practicing mathematics was by themselves. The remaining participants worked with peers in small groups, regularly completed assignments in a Mathematics Lab at the university which facilitates peer-tutoring for lower-division mathematics courses, or consulted tutoring resources. These students described believing that, by working with peers in advanced courses, they benefited from opportunities to share ideas on difficult problems (e.g., proofs) and get feedback on whether their performance on tests was in line with classmates. Charlie even described attending the Mathematics Lab with the express purpose of helping other students in her classes:

> I usually worked in groups. I would go in the Math Lab, and whoever happened to come in, I'd help. And I learn well from teaching…I'd come in [the Math Lab] with the intent of studying, and then people come in and ask for help, or ask me because they recognize me. I think that's why I did as well as I did in a lot of my classes… if you're having to show somebody how to do it, you better figure it out real quick.

Discussion

Several aspects of the research design (e.g., two public regional universities in one state, volunteer interviews of only a few students at one of the universities) and limited scope of the findings (e.g., grade distributions of mathematics majors) suggest the results should be considered exploratory and tied to the context at the two research sites. Nonetheless, the findings suggest several insights into how mathematics programs can help secondary mathematics majors succeed in an undergraduate mathematics programs. The quantitative results suggest that mathematics majors are among the best performing students in mathematics courses, but that many mathematics majors choose not to persist in their undergraduate degree programs. Fewer than half of mathematics majors in the study took more than three mathematics courses during the six year period, and about two-fifths of those who completed seven or more mathematics classes failed at least one course.

From a social cognitive perspective, there are several potential interpretations of high performance paired with low persistence. For instance, (1) mathematics majors may be less likely to persist due to lowered self-efficacy following self-perceptions of poor performance, (2) mathematics majors may have high internal standards for performance that cause them to (accurate or inaccurate) perceive difficulties in advanced mathematics courses as reason to have reduced self-efficacy. Though it's likely that both mechanisms may contribute to switching of mathematics majors, both the cross-tabulations of performance and persistence and interview data both point more to the second possibility. The interview themes particularly highlight a variety of paths secondary mathematics students take toward a bachelor's degree in mathematics, and that those paths may be filled with substantial perceived obstacles to success in mathematics courses, including family concerns, difficulties with instructors, and course content issues. Collectively, the results suggest that even students who have performed exceptionally well in secondary and entry-level college mathematics courses face a relatively large chance of failing at least one course in a mathematics degree program.

One theme that stood out in the analysis of qualitative data was the suggestion that (with the exception of Calculus I), the mathematics courses in which students struggle most tended included the following: difficult content (e.g., abstract concepts, introduction to proof methods), instructors with strict expectations, and tasks in which students are asked to synthesize conceptual elements of content apart from the procedural tasks. With appropriate support for students, we view such courses as providing great opportunities for personal growth in mathematical understanding. Carefully designed courses with titles like Introduction to Proof, Abstract Algebra, and Real Analysis can deeply extend prospective teachers' understanding of secondary content, but they can also serve the important function of recalibrating the beliefs of prospective secondary mathematics teachers who, prior to such courses, may think of mathematics in terms of memorization and assume the main role of mathematics teachers is to lecture (Cooney, Shealy, & Arvold, 1998).

One implication for undergraduate mathematics programs comes from the identification of areas in which undergraduate secondary mathematics majors may seek help in difficult mathematics courses. All interview participants spoke of the importance of obtaining assistance from instructors. When this aid was not seen as productive by the student, he or she may become frustrated and even, as in Katherine's case, decide to change their course of study after just a single negative experience in a mathematics course. Since many mathematics majors have a history of excelling in academics, they have options, and it may be worthwhile for undergraduate mathematics programs to consider formalized structures to facilitate instructor interactions, peer tutoring, and group work among students in advanced mathematics classes. For example, faculty might choose to hold some office hours at an informal Mathematics Lab setting in order to help students in courses such as Calculus II, Real Analysis, Abstract Algebra, or Introduction to Proof.

The findings also suggest strategies faculty might consider in order to help secondary mathematics majors be successful in their mathematics classes. In light of the possibilities that mathematics majors may have experienced secondary mathematics courses emphasizing procedural learning, one potential adjustment university mathematics faculty could make would be to emphasize a standards-based approach (e.g., NCTM, 2000) to teaching mathematics with an emphasis on reasoning and proof in courses at every level. For example, Fukawa-Connelly (2012) has outlined a way in which abstract algebra students can reach deep understanding of proofs by engaging in a student-centered classroom atmosphere in which students present and defend their proofs. Even when formal proofs are not the central focus of an undergraduate mathematics course, instructors can still implement "high discourse communities" (Imm & Stylianou, 2012) in which students' understandings become the focus of class discussions. Not only could students transform the way they learn mathematics, they could break the cycle of procedural-based instruction that makes it difficult for some secondary mathematics teachers to shift to conceptual-based teaching approaches (Cooney, Shealy, & Arvold, 1998).

As exploratory research, a major goal of this study was to suggest avenues for future research into the performance and persistence of prospective secondary teachers in undergraduate mathematics programs. Future plans include broadening the sample pool to include the mathematical experiences of prospective teachers in all STEM fields. Such a study might blend our methodology for analyzing course enrollments with degree attainment outcomes such as those featured in Chen and Weko (2009), which includes quantitative measures of performance and persistence of STEM majors within their respective degree programs. The results from the qualitative strand also point to the value in evaluating effects of interventions designed to support potential secondary teachers with majors from mathematics-intensive fields other than mathematics. One approach would be to investigate effects from intensive faculty-mentoring or peer-mentoring programs on the performance and persistence of prospective secondary mathematics teachers. Several interview participants mentioned the potential value of a Mathematics Lab space staffed primarily by mathematics majors who can offer not only mathematics help, but informal advice about classes and information about grading norms and expectations.

We introduced this study by outlining recommendations from a U.S. government panel (PCAST, 2012) to consider augmenting the "mathematics bottleneck" in the preparation of STEM and secondary mathematics majors by increasing mathematics instruction by non-mathematics faculty. Prior to considering such profound changes, we thought it reasonable to develop a mixed-methods portrait of the experiences of secondary mathematics majors in mathematics courses taught by mathematics faculty. Our study findings suggest that though earning a bachelor's degree in mathematics may include some intense struggles for many prospective secondary mathematics majors, the students can leverage powerful sources of support to earn high letter grades, persist in required courses, and reach their professional goals. Our hope is that these findings can help mathematics programs to assist the many capable mathematics students with an interest in teaching secondary mathematics to succeed in earning a bachelor's degree in mathematics.

References

Bandura, A. (1997). *Self-efficacy: the exercise of control.* New York: W. H. Freeman.

Bryan, T. J. (1999). The conceptual knowledge of preservice secondary mathematics teachers: How well do they know the subject matter they will teach? *Issues in the Undergraduate Mathematics Preparation of School Teachers, 1,* 1-12.

Campbell, N. K., & Hackett, G. (1986). The effects of mathematics task performance on math self-efficacy and task interest. *Journal of Vocational Behavior, 28*, 149-162.

Champion, J. K. (2010). The mathematics self-efficacy and calibration of students in a secondary mathematics teacher preparation program. Dissertation Abstracts International, 71(06).

Chen, P. P. (2002). Mathematics self-efficacy calibration of seventh graders. *Dissertation Abstracts International, 63*(3), 858A. (AAT No. 3047203)

Chen, X., & Weko, T. (2009). Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education. (NCES No. 2009-161). Washington, D.C.: National Center for Education Statistics.

Conference Board of the Mathematical Sciences. (2012). *The Mathematical Education of Teachers II*. Providence, RI & Washington, DC: American Mathematical Society & Mathematical Association of America.

Cooney, T. J., Shealy, B. E., & Arvold, B. (1998). Conceptualization belief structures of secondary preservice mathematics majors. *Journal for Research in Mathematics Education, 29*(3), 306-333.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). Los Angeles: Sage.

DesJardins, S., McCall, B., Ahlburg, D., & Moye, M. (2002). Adding a timing light to the "Tool Box." *Research in Higher Education, 43*(1), 83-114.

Fukawa-Connelly, T. (2012). Classroom sociomathematical norms for proof presentation in undergraduate in abstract algebra. *The Journal of Mathematical Behavior, 31*, 401-416.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42* (2), 371-406.

Imm, K., & Stylianou, D. (2012). Talking mathematically: An analysis of discourse communities. *The Journal of Mathematical Behavior, 31*(1), 130-148.

Ingersoll, R. M. (2003). *Is there really a teacher shortage?* (Document No. R-03-4). University of Washington: Center for the Study of Teaching and Policy.

Kahan, J. A., Cooper, D. A., & Bethea, K. A. (2003). The role of mathematics teachers' content knowledge in their teaching: A framework for research applied to a study of student teachers. *Journal of Mathematics Teacher Education, 6*, 223-252.

Lutzer, D. J. (2002). *Mathematics majors 2002*. Williamsburg, VA: College of William and Mary, Department of Mathematics.

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 CBMS survey. American Mathematical Society.

Ma, X., & Kishor, N. (1997). Attitude toward self, social factors, and achievement in mathematics: A meta-analytic review. *Educational Psychology Review, 9*(2), 89-120.

Monk, D. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*(2), 125-145.

Moore, C. & Shulock, N. (2009). Student progress toward degree completion: Lessons from the research literature. Sacramento, CA: Institute for Higher Education Leadership & Policy.

Morton, B. A., Peltola, P., Hurwitz, M. D., Orlofsky, G. F., Strizek, G. A., & Gruber, K. J. (2008). Education and certification qualifications of departmentalized public high school-level teachers of core subjects: Evidence from the 2003-04 Schools and Staffing Survey (NCES No. 2008-338). Washington, D.C.: National Center for Education Statistics.

No Child Left Behind Act, 20 U.S.C. § 6301 (2001).

Philippou, G., & Christou, C. (1998). The effects of a preparatory mathematics program in changing prospective teachers' attitudes towards mathematics. *Educational Studies in Mathematics. 35*(2), 189-206.

President's Council of Advisors on Science and Technology (PCAST). (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Retrieved from http://www.nctm.org/standards/.

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Working Papers*, Retrieved from http://digitalcommons.ilr.cornell.edu/workingpapers/118.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder: Westview Press.

United States Department of Education Office of Postsecondary Education (2009). The secretary's sixth annual report on teacher quality: A highly qualified teacher in every classroom. Retrieved from http://www2.ed.gov/about/reports/annual/teachprep/t2r6.pdf

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future research. *Review of Educational Research (78)*(4), 751-796.

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal, 37*(1), 215-246.

CHALLENGES WHILE MODELING WHILE USING RIEMANN SUM

Kritika Chhetri
University of Central Arkansas
kchhetri1@uca.edu

Jason Martin
University of Central Arkansas
jasonm@uca.edu

This research focuses on mental challenges that students face and how they resolve these challenges while transitioning from intuitive reasoning to constructing a more formal mathematical structure of Riemann sum while modeling "real life" contexts. A pair of Calculus I students who had just received instruction on definite integral defined using Riemann sums and illustrated as area participated in multiple interview sessions. They were given contextual problems related to Riemann sums but were not informed of this relationship. Our intent was to observe students' transitioning from model of to model for reasoning while modeling these problem situations. Results indicate that students conceived of five major conceptions during their first task and their reasoning about their task from one form of representation became a model for reasoning into another form of representation within a task. Also, their ways of reasoning from their first task served as referential tools to reason about their next task. In this paper we detail those conceptions and their reasoning from their first task that became model for reasoning within and across subsequent tasks.

Keywords: Emergent modeling, Riemann sum, Quantitative reasoning, Definite integral

Introduction and Research Questions

Riemann sums provide a foundation upon which one can understand why definite integrals model various situations found within physics and engineering. Previous research has detailed mental challenges that students face while reasoning about accumulation contexts, and has stressed how students could perform routine procedures for definite integral without being able to explain their reasoning (e.g., Artigue, 1991, Hall, 2010, Orton, 1983; Sealey, 2006). Research that has detailed how students might shift from more intuitive understanding to a more formal understanding has focused on roles of quantitative reasoning (Sealey, 2006; Thompson, 1994) and how that reasoning can support a more conceptually accessible formation of the Fundamental Theorem of Calculus (Thompson & Silverman, 2008). Other research has detailed the importance of conceiving of appropriate structural elements of the Riemann sum within context in order to complete approximation tasks (Sealey & Oehrtman, 2008). But when students come to understand Riemann sums as a model of a particular situation, how does their reasoning about that model influence their reasoning in constructing Riemann sum models of subsequent situations? This research attempts to answer the following questions. (1) What challenges do students face and how do they resolve those challenges as they constitute Riemann sum as a model of a contextual approximation problem? (2) How do students utilize their prior reasoning from their constitution of their Riemann sum model as a model for their reasoning about subsequent problems?

Theoretical Perspectives

Rooted on the theory of Realistic Mathematics Education (Freudenthal, 1973), emergent modeling is an instructional design heuristic where modeling is viewed as an active organizing process where models co-evolve as students reorganize their intuitive reasoning and construct more formal mathematical reasoning (Gravemeijer 2002; Heuvel-Panhuizen, 2003). Consequently, Gravemeijer & Stephan (2010) claim that "formal mathematics is seen as something that grows out of the students' activity […] by way of mathematizing their own informal activities" (p. 148). Gravemeijer & Stephan (2010) classify emergent modeling into four layers; task setting, referential, general, and formal. It needs to be noted that these four layers may or may not appear in a systematic order, as the layers are based on the learner and his learning activities. Task setting involves activities that support the learner to know their problem and understand how to act in a particular situation. Referential activity is also situated to a particular task where learners refer to their task setting activities. As students progress toward the general layer, their activities become more loosely tied to their task setting and they transition from reasoning about a particular problem situation to reasoning about mathematical meanings and the mathematical relationships from their problem situation. As the learner no longer ties their reasoning to a particular task, they generalize their reasoning over related situations and consequently conceive of a new mathematical reality (Gravemeijer & Stephan, 2010; Gravemeijer, 2002, 2004, 2013). "Real life" refers to situations that are experientially real to students, and mathematical reality implies mathematical reasoning that students access intuitively and experience as their reality (Gravemeijer, 2004,2013; Johnson, 2013). The formal layer represents reasoning consistent with a formal mathematical structure that captures all relevant contexts (Gravemeijer & Stephan, 2010).

Models: Models are viewed as more than representations but as holistic organizing activities including a solution strategy. They are a "series of consecutive sub-models that can be described as a cascade of inscriptions or chain of signification" (Gravemeijer, 2004, p. 139). Gravemeijer (2004) refers to this cascade of inscriptions as chain of signification where the central object of the signification involves a sign, "which is made up of a signifier (a name or symbol) and the signified (that which the signifier is referencing, such as the students' activity)" (Johnson, 2013, p. 3). As signifier and signified build robust connections, that then give rise to a new sign, where the new sign then gives rise to another sign, and in this way the chain of signification progresses (Gravemeijer, 2013). These sub-models evolve along students' ways of acting and reasoning about their problem situation and build one after another through an iterative process as seen in Figure 1. These sub-models can be viewed from a micro- and a macro-level when compared to an overarching emergent model of Riemann sum. Sub-models that can be classified as micro-level models, are sub-models that when compared to the overarching emerging model (marco-level model) relate to the emergence of more general reasoning about a particular component of the macro-level model while the macro-level model might still be more directly tied to the problem situation.

In the context of Riemann sum, micro-level sub-models relate to conceiving of quantities and constructing a multiplicative relationship between the quantities of a problem situation, for instance, distance equals the product of velocity at a point times change in time. Whereas macro-level sub-model relates to the idea of summing up individual distances to compute total distance traveled by an object. By actively mathematizing (organizing and symbolizing) their reasoning within multiple representations within a task, for instance, pictorially drawing their situation, conceiving of and writing formulas for their situation, writing contextual sentences about their situation, and graphing their situation, students can develop rich understanding about their situation (Oehrtman, 2008). As they start to conceive of their micro-level sub-models as their new mathematical reality, they transition from 'model of' to 'model for' on a micro-level. Similarly, by computing a series of macro-level sub-models, students transition from 'model of' to ‘model for’ on a macro-level. The micro-level informs the macro-level but a ‘model of’ to ‘model for’ transition at the macro-level need not imply this transition at the macro-level. Through the interplay between the two levels and the ‘model of’ ‘model for’ transitions, students construct series of consecutive sub-models in a spiral fashion (Figure 1) by conceiving of, organizing, and reasoning about multiple representations of their problem situation.

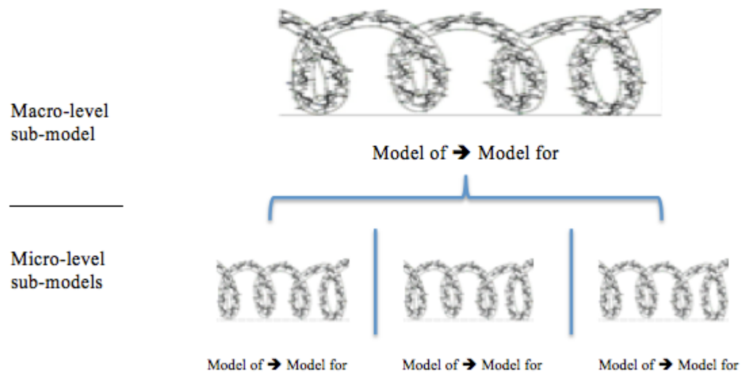


Figure 1: Series of evolving models in an iterative process.

Model of- Model for: *Model of* is the starting phase of emergent modeling where learners consider a model to be context-specific and employ informal solution strategies. As they employ their informal reasoning about their emerging “models of” and start to identify commonalities amongst different situations, their model starts to change characters, and slowly comes to life of its own, with its own identity. Model of involves activities in the task setting and referential layer. As students move from referential to general, they transition from model of to model for. Model for is the latter phase of emergent modeling where learners shift from thinking about the problem situation of the model to reasoning about it mathematically. “The model changes character, it becomes an entity of its own, and as such it can function as a *model for* more formal mathematical reasoning” (emphasis in original, Gravemeijer, 2002, p.2). Model of- model for is a mental shift, which continually occurs within students’ mind and evolves along their ways of organizing and symbolizing. During this shift, learners gradually shift from thinking about a problem situation to reasoning about mathematical relationship of the quantities of their problem situation to organizing and symbolizing those emerging reasoning to construct a new generalizable formal mathematics (Gravemeijer & Stephan, 2010; Gravemeijer, 2013). For example, in the context of Riemann sum, this shift involves a transition from viewing a summation of a multiplicative structure of a problem situation to being able to find commonality across different problem situations and then generalize those reasoning about those contexts to reason about other similar contexts related to Riemann sum.

Quantitative reasoning: “Objects are constructions that a person takes as given and that qualities of an object are imbued by the subject conceiving it” (Thompson, 1994). Smith and Thompson (2007) say that quantities are “ measurable attributes of objects or phenomena” (as cited in, Larson, 2010, p. 112). Quantitative reasoning provides a means of modeling where students conceive of quantities, construct relationships between quantities, and meaningfully operate on those quantities that can support the construction of further quantities as one reasons with and about the problem situation (Larson, 2010; Thompson, 2011). When a conceived quantity is specifically attached to an attribute of a problem situation, representing of this quantity would indicate model of reasoning, but as one reasons about this quantity within a quantitative structure without referring to a problem situation, that reasoning emerges as a model for their reasoning about the mathematics.

In this study we use the approach of emergent modeling from RME to observe how students construct and reason about their micro-level sub-models while organizing and symbolizing within multiple ways of representation. For instance, what/how does their pictorial reasoning supplement their numeric, algebraic, verbal and graphical organizing and symbolizing? Also, how do they construct their micro-level sub-models while reasoning about their problem situation, and what/how do they conceive of their emerging micro-level sub-models as their new mathematical reality, and how do they employ their newly emerged mathematical reality into constructing/generalizing their macro-level sub-models.

Method

Ten interview sessions (50-148 minutes) were conducted with two volunteer Calculus I students, Sam and Chris (pseudonyms), who had been introduced to the definite integral through Riemann sums illustrated as area under the curve using Stewart (2008). These students had not been exposed to Riemann sum as modeling other situations besides area. During the interviews, Sam and Chris were given three approximation tasks related to Riemann sums, out of which two emphasized finding under and overestimates to total distance traveled based off of a table containing velocities and a velocity function, respectively (Figure 2). The third task related to pressure on a dam, but this paper will focus on only the first two tasks since analysis of the third task is ongoing. While the first two tasks allows us to observe the emergence of micro-level sub-models while reasoning specifically about distance, rate, and time relationships, the third task allows us to see how the modeling activities from task 1 and 2, facilitated students to conceive of and organize emerging model of Riemann sum at the macro-level.

For the first two tasks, additional subtasks included drawing pictures of the actual situation, finding and illustrating error bounds, and graphing. The first two tasks involved a lot of scaffolding to facilitate students' transition from informal to formal mathematics as they worked in a pair. However, scaffolding was reduced from third task to observe how they would progress with reduced help. Sessions were audio and videotaped to analyze how students modeled their problem situations. Models were identified based on students' reasoning as exemplified by their representations and verbal utterances. When students directly related their reasoning to the -

Task 1: The table below shows the velocity of a car travelling from Conway to Little Rock. In this activity you will approximate the distance travelled by the car during the first 10 seconds of the car entering the southbound I-40 ramp.

T(s)	0	2	4	6	8	10
V(ft/s)	0	21	34	44	51	56

Task 2: NASA's Q36 Robotic Lunar Rover can travel up to 3 hours on a single charge and has a range of 1.6 miles. After t hours of traveling, its speed in miles per hour is given by the function $v(t) = sin\sqrt{9-t^2}$. In this activity you will approximate the distance travelled by the Lunar Rover in the first two hours.

Figure 2. First two teaching experiment tasks.

problem situation, this was viewed as model of reasoning. Prior patterns of reasoning and representing when applied to a current problem situation were viewed as indicators of potential model for reasoning. The sessions concluded with a generalizing activity where students were asked to expand and extend their reasoning about their three tasks into a general formula, which could encompass similar situations. This final task was designed to provide insight into the more general and formal layer of emergent modeling as it related to the macro-level model of an emerging Riemann sum model.

Results

The results reported here will focus on the students' emerging model of Task 1 (Table 1) and reasoning about Task 1 that reappeared in Task 2 to suggest at least the referential layer of reasoning and a potential movement toward the general layer of reasoning.

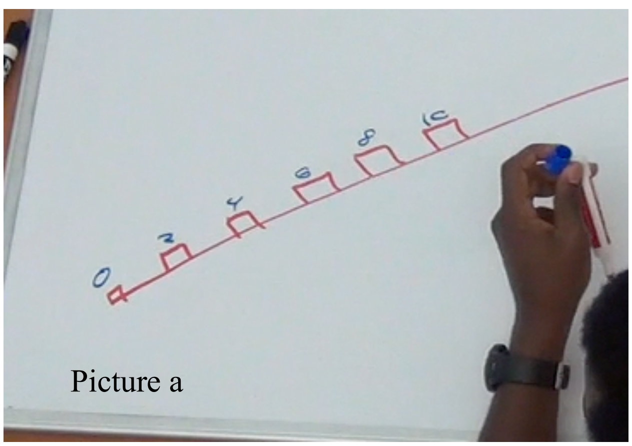
Table 1.
Distinct Conceptions During Task 1.

Level of Conception	Conception	Description of reasoning
Micro-Level	DRT 1: Distance changes as time changes	Omitted explicit detail to amounts of change in velocity. Pictorially represented as a vehicle with constant amounts of changes in distance per two-second intervals.
	DRT 2: Distance is change in velocity × change in time	Initially supported by their reasoning about amounts of change in distance vary because of changing velocities. Pictorially represented as a vehicle with decreasing amounts of changes in distance per two second intervals which became a model for distance as $d = \Delta V \cdot \Delta t$.
	DRT 3: Distance is constant velocity × change in time	Initially only conceived for a vehicle traveling at constant velocities. After adjusting their picture to model a vehicle with increasing amounts of changes in distance and after "supposing" their vehicle as traveling at constant velocities was this conception applied to their context. Formulaically represented as $d = V_p \Delta t$.
Macro-Level	Total 1: Total distance approximated by adding up distances are underestimates or inconclusive.	Adding up amounts of change in distances approximates total distance. Coordinated with DRT 2 and then DRT 3. With DRT 3 it was initially represented as $\sum_{p=0}^{5} V_P \Delta t$. For Sam, this sum was an underestimate because the sum would increase towards the exact total distance traveled as more data points were added. For Chris, this sum was inconclusive because the data table did not reveal what happened between data points.
	Total 2: Total distance approximated by adding up using max. and min. velocities.	Coordinated with DRT 3. They conceived of maximum and minimum velocities over a time-interval as approximations to varying velocity over that interval. Underestimates and overestimates were represented by $\sum_{p=0}^{4} V_P \Delta t$ and $\sum_{p=1}^{5} V_P \Delta t$, respectively.

Note. DRT = Distance, Rate, and Time relationship.

During the first session, Sam and Chris struggled with their conceptions of the relationship between distance, rate, and time (referred to as DRT conceptions). Since our intent was to observe their evolving formation of a Riemann sum model, their evolving DRT model can be viewed as a micro-level model when compared to the overarching emerging Riemann sum model for total distance traveled.

Conceptions of DRT in Task 1: Initially, Sam and Chris realized that the varying velocities and the finite amount of data caused problems with easily completing Task 1. Reasoning from the provided table, their first conception of a distance/rate/time relationship (DRT 1[1]) was modeled as a picture containing snapshots of a car equally distanced between every two seconds (Figure 3, Picture a). After the facilitator prompted them to be "picky" with their picture, they attended to varying amounts of change in distance between snapshots, and represented this conception pictorially with increasing changes in distance between every two-second snapshot (Figure 3, Picture b) and formulaically as "$d = \Delta V{\cdot}\Delta t$" (DRT 2). Then they calculated distance traveled by the car over each interval. For the first and second intervals they came up with 42 ft and 26 ft, respectively. At this point, their picture (as seen in Figure 3, Picture b) served as a 'model for' their calculations that appear in Figure 4. They reasoned that 42 ft was greater than 26 ft as demonstrated on their picture, and they did not have any intellectual conflict to rethink about their picture, quantities or formula.



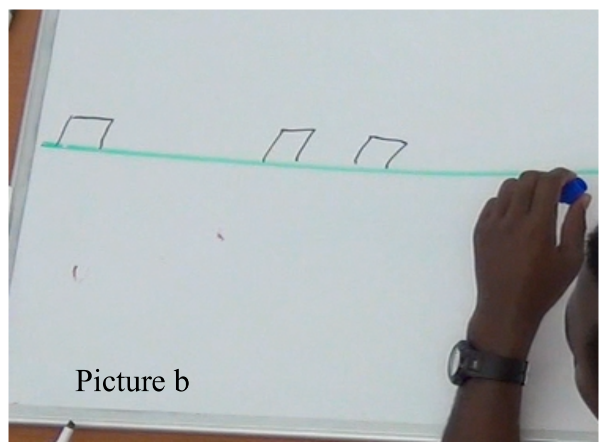


Figure 3. Two pictures showing locations of the car every two seconds.

Only after they were prompted to think about a "real life" situation of a car merging onto an interstate, did they adjust their picture to indicate increasing distances between snapshots. Responding to "real life" situation, Sam instantly drew a picture of a moving car (Figure 5, Picture a) and said, "its where it starts, would be the next place and the next place will be little bit further and further and further." Initially, Chris did not recognize the difference between their previous picture and Sam's new picture, but as Sam explained to him about the amounts of change in distance caused by the amounts of change in velocity, Chris adopted Sam's new picture. Together, Sam and Chris reorganized their pictures (Figure 5, Picture b) and reasoned that greater velocity would yield greater distance.

[1] Descriptions for DRT 1, DRT 2, DRT 3, Total 1, and Total 2 can be found in Table 1

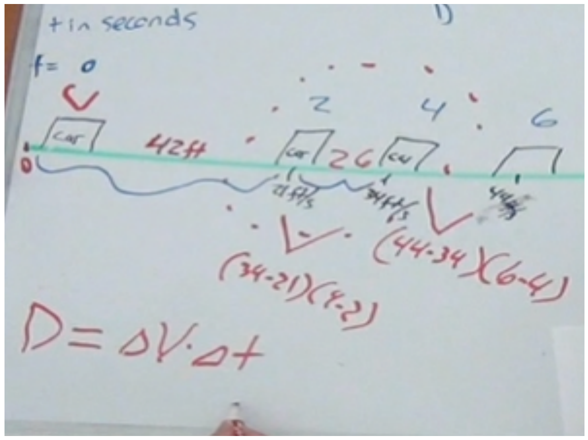
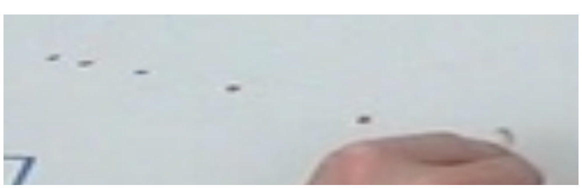


Figure 4. Representations of DRT 2.

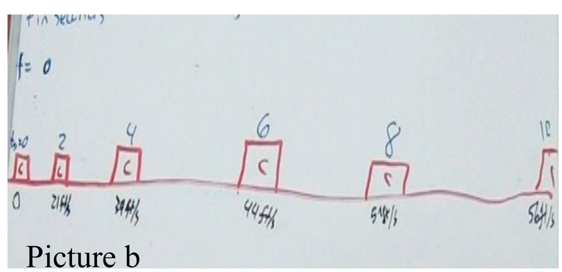
Although by this moment they had indicated that adding up individual distances would provide approximations for total distance (Total 1) they still struggled to symbolize their emerging reasoning about distance. At first, DRT 3 appeared in response to an additional facilitator question concerning another situation in which a car traveled at 70 mph for two hours and 80 mph for one more hour. Though they concluded that the car in the other situation traveled 220 miles, DRT 2 persisted in their reasoning about the car with varying velocities in Task 1. Once they realized that their formula $d = \Delta V{\cdot}\Delta t$ for the other situation yielded a conflicting answer when applied to the additional question did they rethink DRT 2. Meanwhile, their picture and their calculated distance supported them in reasoning about velocity at a point and constructing DRT 3. They symbolized DRT 3 as $d = V_p\Delta t$, and verbalized that distance would be the product of velocity at a point times change in time.

Conceptions of Total Distance in Task 1: After conceiving DRT 3, they summed up their individual distances over an interval to 412 ft But when asked if that was an underestimate, overestimate or neither, Sam initially responded that, "412 ft is an under approximation of the distance traveled because we [can] have more data points." However, Chris mentioned that he was not sure if 412 ft was under, over or neither. Chris stated:

> [The driver] didn't come from 21 to 34 immediately. […] Because maybe before he get to this 34 ft/sec, there was an intermediate velocity that we don't know here. Which might be bigger or smaller [than 34]. So that's why I can- I cannot say this is the underestimation.



Picture a



Picture b

Figure 5. Graphical representations of DRT 3.

Attempting to calculate error bound, they grappled with finding both under and overestimates for total distance. After three hours since starting this task, once they had conceived of the roles of maximum and minimum velocities as approximations for varying velocities over two-second intervals, they coordinated DRT 3 with total distance and were able to find both under and overestimates for total distance (Total 2). Having been prompted by the interviewer to reason about 21ft/s as a minimum or maximum velocity over the first two time intervals, they concluded that 21 ft/s was the maximum velocity over the 0-2 second time interval and a minimum velocity over the 2-4 second time interval provided that the car was speeding up. As they conceived of maximum and minimum velocity for an interval, they employed that reasoning about maximum and minimum velocities over the remaining time intervals and articulated that 412 ft would be an overestimate. Sam concluded the following for the pair:

> The displacement of the car in 10 seconds is an overestimate because if we consider 56 feet per second for the total 10 seconds is a maximum overestimate of the displacement then it follows that for each point the maximum velocity observed taken for total velocity at that point leads to a maximum overestimation of displacement at that point.

It was not until they had to find the error bound that they felt the need to employ similar reasoning about minimum velocity over a time interval.

Given that their teacher had already modeled Riemann sum as the area under the curve, we expected that graphing distance would be very salient. However, the students were not able to immediately graph their situations represented as area under a velocity function. Instead, how they reasoned about and constructed their picture served them as ‘model for’ finding distance on their graph. As they proceeded to label and compute their distance, they attended to secant lines of their curve (Figure 6) as their distance travelled by the car instead of area under the curve. The way their distances appeared in their picture (Figure 5, Picture b) became the signifier, which signified to them that those distances would appear like a straight line in their graph as well. In this sense, the picture became a ‘model for’ their reasoning concerning their graph. Unfortunately, these straight-line distances from their picture were initially conceived of as secant lines on their graph and not as area.

Excerpt 1:

I: Where is delta t times velocity? How would that look over here [points to their graph]?

Sam: This, this line right here [puts his pen on the secant line to their graph between time interval 2-4s] between the two points.

I: Why so?

Sam: Because, like I was saying, the other side of the triangle, here is our change in velocity [points to the hypotenuse of his right angled (almost) triangle] between the two points and there is our delta t.

Chris: If you consider the expression, the displacement will be these [draws more individual secant lines], all these line[s] it will be the total displacement [stretches his hand over the secant line over the entire change in time].

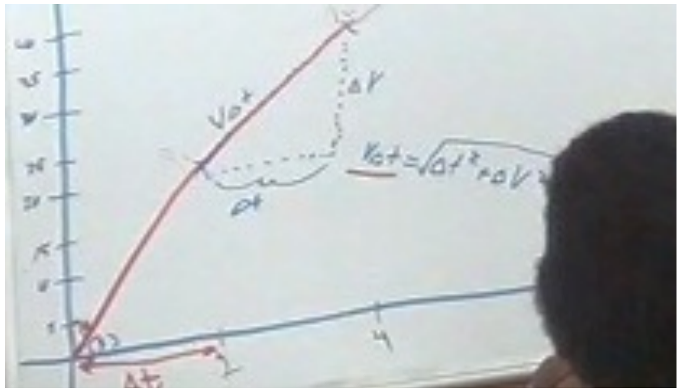


Figure 6: Distance as secant lines.

Note how in Figure 6, the students had even labeled the distance that the car traveled between 2 and 4 seconds as $V\Delta t = \sqrt{\Delta t^2 + \Delta V^2}$. They subsequently acknowledged that $\sqrt{\Delta t^2 + \Delta V^2}$ did not yield the correct units and Sam dropped his notion that distance would appear like straight lines. Even though they wrestled for forty minutes to resolve this issue and incorporate their prior reasoning about distance on their graph, they remained unable to graphically reproduce 412 ft as an overestimate, 300 ft as an underestimate, or any other numerical approximation. During this time after Sam had dropped his notion of distance as straight lines, he did not represent distance's multiplicative relationship with velocity and change in time as area. After the facilitator stepped in and merely drew one approximating rectangle over the first time interval, Sam immediately conceived of approximations to distance through rectangular areas. Sam then explained to Chris, who then accepted Sam's explanation. They subsequently identified their underestimate as the sum of areas of rectangles below the curve and overestimate as the sum of areas of rectangles above the curve. As they resolved their issue of representing total distance as the sum of areas of rectangles, they coordinated their prior reasoning about their picture, table and algebraic expression about error and error bound into their graph and reasoned that the bigger the number of their rectangles, the better their approximations would be.

Later Sam compared getting an exact distance to a perfect video, "We have an infinite number of snapshots, […] a solid image of what- We have a video, a perfect video where there is no frames or anything like that, an ideal video." They finished with representations seen in Figure 7.

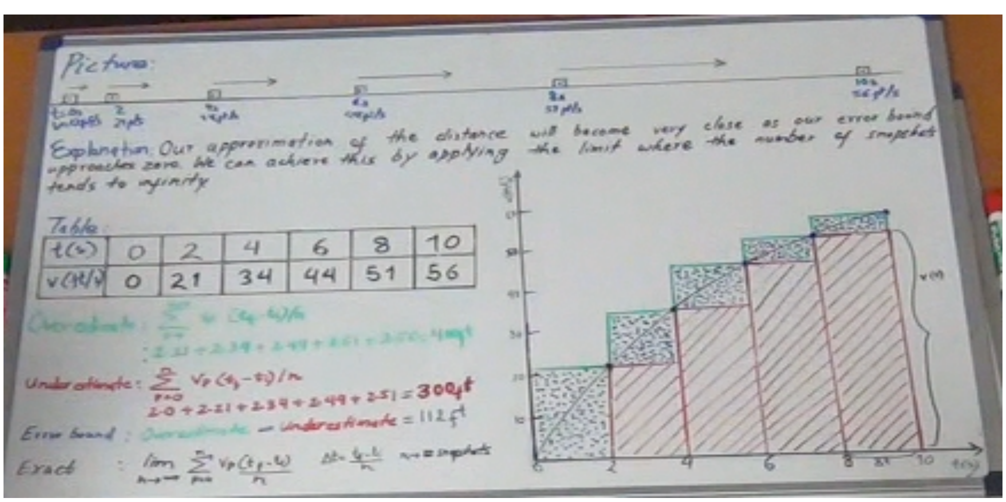
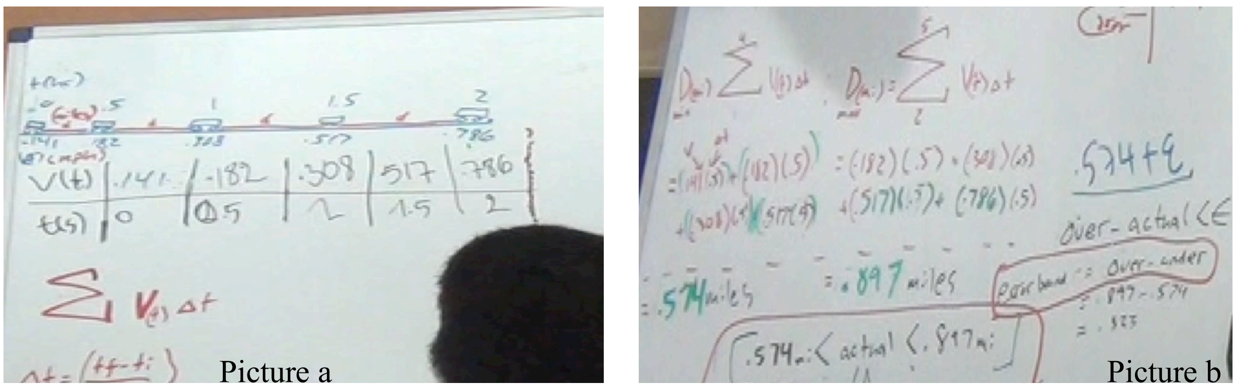


Figure 7. Chris and Sam's multiple representations of Task 1.

Task 2: Immediately after being given Task 2, Sam asked Chris, “Don’t you think the picture looks the same like last time?” He then drew a picture with snapshots of the rover at every half hour interval (Figure 8, top of Picture a). In response to Sam’s question, Chris said, “But, do you know this rover is accelerating or decelerating or and anything…lets make a table.” In order to understand the changing velocity of the rover, they constructed a table of values (Figure 8, Picture a) similar to the given table in Task 1. It can be inferred that reasoning through picture became a ‘model for’ reasoning for Sam while reasoning through table became a ‘model for’ reasoning for Chris. Once their table was constructed, they explicitly quantified amounts of change and coordinated that with labels added to their picture. After being asked, “Where is distance?” they noted distances between snapshots on their picture as they did on their previous task. Sam referred back to Task 1 as he said, “um-hm. Can do the same which we did last time. Sum up velocity, we will say velocity at that time times the change in time.” Both of them agreed that finding total distance was, “the same as what we did last time.” Subsequently, they represented total distance as $\sum V_{(t)}\Delta t$ (Figure 8, bottom of Picture a).



Picture a Picture b

Figure 8: Sam and Chris' task 2 representations.

Although imprecise, their representation in Figure 8 captures the multiplicative structure between particular velocities and amounts of change in time within a summation. More than the symbolization, it was their way of symbolizing from their Task 1 that became a referent for their Task 2 symbolizations. For instance, their ways of reasoning about maximum and minimum velocity over a time interval in their previous Task 1 allowed them to attend to and reason about maximum and minimum velocity in Task 2. For example, when prompted about total distance, Chris replied, “We have the same issue as last time where we have to consider the maximum velocity or the minimum velocity.” In the process of constructing a numerical approximation to total distance travelled, they employed their prior reasoning concerning maximum and minimum velocity over an interval and coordinated their DRT 3 and Total 2 from their Task 1 to calculate under and overestimates, eventually represented as summations with appropriate adjustments to the starting values of the index (Figure 8, Picture b). Even within their Task 2, their ways of reasoning about their picture and table again served as ‘models for’ constructing their algebraic expression for their situation. However, if we look across tasks, it can be said that how they constructed their algebraic expressions for under and overestimates during Task 1 served as referential tools in constructing similar expressions for Task 2. In the same way, how they reasoned algebraically in Task 1 served them in constructing expressions for Task 2, and their ways of constructing and reasoning about their Task 1 graph helped them in constructing their Task 2 graph (Figure 8). As soon as they labeled their graph, Sam exclaimed, “Ok now, so [we] gotta do rectangles!" and both of them continued to draw under and overestimate rectangles as representations for under and overestimates of distance traveled. Sam further contrasted how quickly they were able to represent distance traveled, "we got a minimum one and it took so long to find last time. This is our minimum [points to rectangles below the curve]."

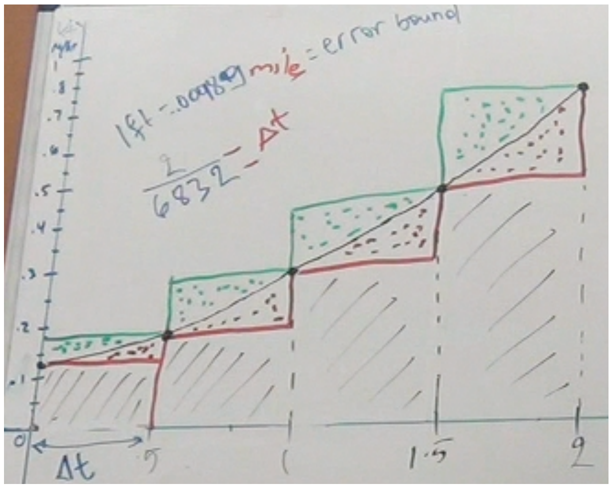


Figure 8: Graphical depiction of task 2.

From approximations to exact distance, Sam again brought up his analogy with video, "We go from having pictures, to a flip book, to a video, to like one true continuous string where there is no frame rate." After finishing Task 1 and 2, they reflected on their task as follows:

Excerpt 2:
Sam: So the first time we just had data and we kinda of wondered about we just had data points. The second time, we had an equation and we figured out our own points based on the equation.
Chris: um um
Sam: and then we refined our, our approximation, or our approximations during the second time.
Chris: all of these summarize how you can make an error smaller by increasing the number of snapshots[...]as we increase the number of snapshots we tending to get the exact displacement so, that's what both of them summarize [points to both Task 1 and 2]

For the final generalizing activity, in addition to writing a generalizable formula, students also expressed that they could calculate force (mass and acceleration) and current in the same way that they calculated distance and force in the teaching experiment.

Discussion

For this study we asked, (1) What challenges do students face and how do they resolve those challenges as they constitute Riemann sum as a model of a contextual approximation problem? (2) How do students utilize their prior reasoning from their constitution of their Riemann sum model as a model for their reasoning about subsequent problems?

In the results we detailed the challenges that Sam and Chris faced in not only conceiving of a Riemann sum structure but in conceiving of a relationship between distance, rate, and time consistent with finding estimates for total distance traveled. How these challenges were resolved were captured by the movements from one conception to the next conception as presented in Table 1. We observe that the challenges encountered during Task 1 were not easily overcome by Sam and Chris but their engagement of these challenges supported them in forming patterns of reasoning for more effectively modeling Task 2. We note that it was not merely the end results of Task 1, but elements of their reasoning that went behind creating those end results, including a solution strategy, which served as a model for subsequent reasoning during Task 2. By "elements of reasoning" we mean reasoning about components related to a Riemann sum context that can be viewed at either the micro-level (e.g. distance, rate, and time relationship) or the macro-level (emerging Riemann sum model used to capture total distance).

At the micro-level, Sam and Chris' transition from 'model of' to 'model for' was apparent within a task in their ways of reasoning, organizing, and symbolizing as they progressed from pictures to graph. For example, during Task 1, Sam and Chris symbolized secant lines as their distances by referring back to those straight lines from their pictures and, in this sense, those straight lines in their picture became a 'model for' reasoning about their graph. Their sub-models of the distance, rate, and time relationship they constructed in Task 1 served as models for constructing a similar sub-model of distance, rate and time relationship for Task 2. Similarly, they also carried over their reasoning of adding up individual distances to total distance in Task 1 and later to total distance in Task 2. The use of these micro-level sub-models within and across tasks suggests that reasoning about these components transitioned at times from 'model of' to 'model for' type reasoning.

From a macro-level perspective, their ways of reasoning about distance, rate and time for a car of increasing velocity pictorially, algebraically, numerically and graphically in Task 1 served as referential tools for constructing a total distance in Task 2. In this sense, their 'model of' summing up individual distances using their DRT relationship to find total distance in Task 1 served as a referent for reasoning about and organizing a model of the summation in Task 2. In addition, their picture, graph, table, and formulaic expressions from Task 1 served as reference points for them to make connections between their two tasks as they conceived of, represented, and related relevant quantities. For example, before they firmly committed to using their reasoning from Task 1 applied to Task 2, their pictures and tables supported their conceiving of varying velocities, amounts of change in time, amounts of change in distance, and in relating these quantities while building connections across the tasks. As these connections became more apparent, the students progressed in constructing appropriate Riemann sum approximations. Furthermore, in Task 1 they had to conceive of pertinent roles for minimum and maximum velocities for under and overestimates, and relate those to a notion of summing up distances to obtain Riemann sums for under and overestimates. This supported Sam and Chris' conceptions of minimum and maximum velocities, which emerged after three hours of work during Task 1, to be readily represented as estimates for total distance for Task 2 after only thirty-two minutes.

Limitations and Questions

We acknowledge that our work with one pair of students does not necessarily generalize to others. Moreover, at the macro-level not much can be concluded about the transition from 'model of' to 'model for' as most reasoning cannot be well distinguished between the referential and the general layers of reasoning. Even with a third task, distinguishing well between these two layers will be difficult. We also note that since the students were exposed to Riemann sums, they were not directly reinventing Riemann sum symbolizations but were reinventing a multiplicative structure within contexts and constructing relationships between this structure and an emerging additive structure.

Our presentation questions were: How can we design tasks to better capture students' modeling activities and their transition from model of to model for in the context of definite integral and Riemann sum within a research context? For students who have not been exposed to Riemann sums, how can we modify our tasks to generate an intellectual need for these sums and subsequently support these students in constructing a Riemann sum? How might activities be effectively scaffolded to support the model of / model for transition in a classroom? We highlight some important feedback concerning these questions that we received during the RUME conference. To better support students in coming to reason about the distance, rate, and time relationship while leveraging ideas that emerge from student reasoning, a pre task might be designed to support students in coming to understand how distance accumulates over varying amounts of time and velocities. A subsequent non-scaffolded novel Task 4 might also be added after the generalizing activities following Task 3 to illuminate 'model for' reasoning at the macro-level. These changes to the tasks represent potential future studies that can better address our research questions.

References

Artigue, M. (1991). Analysis. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 167-198). Boston: Kluwer.

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht, Holland: D. Reidel

Gravemeijer, K. (2002). Emergent modeling as the basis for an instructional sequence on data analysis. In *Proceedings of the 6th International Conference on Teaching Statistics*. Voorburg, the Netherlands.

Gravemeijer, K. (2004). Emergent modeling as a precursor to mathematical modeling. In H. W. Henn & W. Blum (Eds.), ICMI study 14: Applications and modelling in mathematics education, pre-conference volume (pp. 97-102). Dortmund: Universitat Dortmund.

Gravemeijer, K. (2013). How to support students in constructing more formal mathematics. Presented at *The 16th Conference on Research in Undergraduate Mathematics Education*. Denver, CO: University of Northern Colorado.

Gravemeijer, K., & Stephan, M. (2010). Emergent models as an instructional design heuristic. In Gravemeijer et. al. (Ed.), *Symbolizing, modeling, and instruction design* (pp. 145-169). Mahwah, NJ: Erlbaum and Associates.

Hall, Jr., W. (2010). Student misconceptions of the language of calculus: Definite and indefinite integrals. *In Proceedings of the 13th Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC: North Carolina State University.

Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics, 54*(1), 9-35.

Johnson, E. (2013). Implication of Realistic Mathematics Education for analyzing student learning. *In Proceedings of the 16th Conference on Research in Undergraduate Mathematics Education*. Denver, CO: University of Northern Colorado.

Larson, C. (2010). Modeling and quantitative reasoning: The summer jobs problem. In R. Lesh, C. R. Haines, P. L.Galbraith & A. Hurford (Eds.), *Modeling Students' Mathematical Modeling Competencies: ICTMA 13* (pp. 111-118). New York, NY: Springer.

Oehrtman, M. (2008). Layers of Abstraction: Theory and design for the instruction of limit concepts. In M. P. Carlson, C. Rasmussen (Eds.). Making the Connection: Research and Teaching in Undergraduate Mathematics Education (Vol. 73, pp. 65-80). Washington, DC: Mathematical Association of America.

Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics, 14*(1), 1-18.

Sealey, V. (2006). Definite integrals, Riemann sums, and area under a curve: What is necessary and sufficient? *In Proceedings of the 28th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Merida, Mexico: Universidad Pedagogica Nacional.

Sealey, V., & Oehrtman, M. (2007). Calculus students' assimilation of the Riemann integral. *In Proceedings of the 10th Conference on Research in Undergraduate Mathematics Education*. San Diego, CA: San Diego State University.

Stewart, J. (2008). *Calculus: Early Transcendentals* (6th ed.). Belmont, CA: Thomson Learning, Inc.

Thompson, P. (1994). Images of rate and operational understanding of the Fundamental Theorem of Calculus. *Educational Studies in Mathematics*, 26(2-3), 229-274.

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

Thompson, P., & Silverman, J. (2008). The concept of accumulation in calculus. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (Vol. 73, pp. 117-131). Washington, DC: Mathematical Association of America.

EXPLICATION AS A LENS FOR THE FORMALIZATION OF MATHEMATICAL THEORY THROUGH GUIDED REINVENTION

Realistic Mathematics Instruction supports students' formalization of their mathematical activity through guided reinvention. To operationalize "formalization" in a proof-oriented instructional context, I adapt Sjogren's (2010) claim that formal proof is an explication (Carnap, 1950) of informal proof. Explication is the process of replacing unscientific concepts with scientific ones. I use Carnap's criteria for successful explication to demonstrate how each element of mathematical theory (definitions, axioms, theorems, proofs) explicates its less formal correlate. I provide examples of students' proving activity in an axiomatic geometry course to motivate the need for explication, meaning that students should see formal theory as a precise expression of their less formal understandings. I conjecture that students who understand formal theory as an explication will better coordinate semantic and syntactic reasoning toward proof. I provide supporting evidence for this claim from a teaching experiment in which students reinvented axioms and definitions of planar geometry.

Key words: Explication, proof, axiomatic geometry, realistic mathematics education

Terms like "formal" and "rigorous" are ever-present in discussions of proof and proving, but often lack clear definitions or even descriptions in practice. Sjogren (2010) claimed that formal proof could be understood as the *explication* (Carnap, 1950) of informal proof. Carnap defined explication as the process of replacing or supplanting informal or unscientific concepts with formal or scientific ones. While not every mathematics student perceives that advanced mathematical theory expresses, generalizes, or formalizes their prior mathematical understandings, this is an express goal in the Realistic Mathematics Education tradition (Freudenthal, 1973; Gravemeijer, 1994). Gravemeijer's four stages of mathematical activity describe how students segue less formal or more situational mathematical activity into more formal understandings through abstraction and generalization. Thus there appears room for utilizing Sjogren's version of explication in conjunction with RME's practice of guided reinvention. In this paper, I explore how Carnap's criteria for explication provide supplementary guiding criteria for analyzing and fostering students' proof-oriented activity as they are engaged in formalizing their pre-formal understandings through guided reinvention.

These guiding criteria, which I use both for analysis and instructional design, developed out of my research on students' learning in an undergraduate, axiomatic geometry course. This is a particularly appropriate context for harnessing students' less formal activity because:

1. students enter undergraduate geometry courses with rich spatial, experiential, and less formal knowledge (i.e. - pre-formal conceptions in need of explication),
2. the course which housed these investigations defines and formalizes very basic experiential concepts such as distance, spatial arrangement, rays, and lines, and
3. while the modern axiomatic tradition emphasizes translating geometric proofs into some purely syntactic calculus native to the axiomatic system, students need semantic insight (Weber & Alcock, 2009) to guide geometric proof production.

I endorse reinvention/explication as an appropriate response to the third observation because it guides students to create formal elements of theory (axioms, definitions, theorems, proofs) as expressions of more intuitive concepts, classes, and ways of reasoning. Whereas previous research tends to dichotomize semantic and syntactic proof production (Alcock & Inglis, 2008), adopting an explication lens problematizes students' ability to coordinate the two.

My central hypothesis is that students engaged in reinvention of elements of formal theory will have better abilities to realize the benefits of both semantic and syntactic reasoning for successful proof production. By blending explication and reinvention, I extend Sjogren's (2010) claim about formal and less formal proof. I argue that in a reinvention-oriented instructional setting, definitions, axioms, theorems, and proofs can all serve as explications of their respective less formal correlates in student thought and activity. In what follows, I elaborate this hypothesis by identifying these "less formal correlates" based on previous literature and examples drawn from my investigations in axiomatic geometry learning. I provide two comparative cases of students' proof activity featuring students taught with and without a reinvention orientation.

A pilot study in axiomatic geometry

To clarify the use of the theoretical tools in this paper, I shall first orient the reader to the mathematics being taught in the axiomatic geometry course and present a proving episode from a pilot study. This episode provides a representative example for the challenges students face in producing proofs in axiomatic geometry motivating the need for explication.

The body of geometric theory

The textbook Blau (2008) presents a simultaneous (or "neutral") axiomatization of the three classical examples of absolute planes: Euclidean ($\mathbb{E}$), Spherical ($\mathbb{S}$), and Hyperbolic ($\mathbb{H}$). The book includes 21 axioms introduced over nine chapters from which a characterization theorem is proven stating that any *absolute plane* (satisfying all 21 axioms) must be of one of the three types. However, since the axioms are introduced in stages, the set of example planes begins very large (including finite geometries) before being narrowed by the introduction of new axioms. The other example planes serve as counterexamples for possible theorems ("Axioms 1-$n \Rightarrow$ Axiom (n+1)" or "Axioms 1-$n \Rightarrow$ Theorem X") and motivate the introduction of new axioms.

According to Blau (ibid.), any *plane* as a collection (**P**, **L**, d, [ω], μ) where **P** is the set of points, **L** is the set of subsets of **P** called lines, d:**P**x**P** $\rightarrow \mathbb{R}$ a real-valued distance function (AB=d(A,B)), ω the sup of the set of all distances (called the diameter of the plane), and μ the angular distance function. Table 1 presents how other geometric concepts are defined from these components of any plane. These definitions exemplify how semantic reasoning about the example planes is necessary for understanding the body of theory in this course, since the condition that segments are only defined for endpoints A and B such that AB< ω is trivial on the Euclidean and Hyperbolic planes where $\omega = \infty$. Blau (2008) includes the condition to accommodate the Spherical plane where the set $\overline{\mathrm{AB}} = \{\mathrm{A,B}\} \cup \left\{\mathrm{X|A\text{-}X\text{-}B}\right\}$ includes the entire Sphere whenever A and B are antipodes (such as the North and South Poles).

Concept	Intuitive meaning	Notation	Definition (Blau, 2008)
Point B lies between points A and C (called a "betweenness relation")	B lies on the linear path from A to C or B is spatially between A and C	A-B-C	Points A, B, C are distinct and collinear such that AB+BC=AC
Line segment with endpoints A and B	The section of a line between and including A and B	$\overline{\mathrm{AB}}$	Given that 0<AB< ω, $\overline{\mathrm{AB}} = \{\mathrm{A,B}\} \cup \{\mathrm{X|A\text{-}X\text{-}B}\}$
The ray with endpoint A through point B	Half-line or all the points along a path beginning at A in the direction of B	$\overrightarrow{\mathrm{AB}}$	Given that 0<AB< ω, $\overrightarrow{\mathrm{AB}} = \{\mathrm{A,B}\} \cup \{\mathrm{X|A\text{-}X\text{-}B}\} \cup \{\mathrm{X|A\text{-}B\text{-}X}\}$
Four point betweenness	Four collinear points sequentially arranged, entailing four betweenness relations	A-B-C-D	A-B-C, A-B-D, A-C-D, and B-C-D

Table 1. Blau's (2008) definitions for basic geometric properties and objects in the plane.

Explication in axiomatic geometry

Carnap was a proponent of logical positivism in the sense that he believed knowledge advanced toward objective truth via the expansion of scientific analysis. Explication to him was an historical phenomenon by which unscientific concepts were replaced by concepts ready for scientific treatment. He claimed a formal concept explicates a less formal one whenever it has *similarity*, *exactness*, and *fruitfulness*. Carnap (1950) provides the example of how temperature explicates warmness. Warmness is an experiential concept that is relative to the observer (one room may be warm coming in from a winter day and not warm stepping out of a bath). Temperature is *similar* to warmness in that most experiences of "warm" can be expressed and explained in terms of "temperature." Temperature differs in that it is an *exact* concept embedding warmness in a scientific body of theory: numerical measurement or quantity. Finally, temperature explicates warmness because it is *fruitful* for the construction of further scientific theory such as the Ideal Gas Law (PV=nRT).

My use of *explication* differs from Carnap's (1950) because I am interested in conceptual shifts in the understanding of individual students, which he explicitly distances himself from calling it "psychologism" (p. 41). As such, I refer to the processes I analyze in this paper as *psychological explication*, the process by which an individual supplants or corresponds an intuitive or less formal concept with a formalizable concept appropriate for the construction of mathematical proof. I claim that the definitions in Table 1 represent explications of the associated intuitive concepts. For instance, the definition of ray is *similar* because the defined set coincides with standard Euclidean rays. Blau (2008) defines spherical geometry where every point on the sphere is distinct, lines are great circles, and distances correspond to the length of the minor arc between points. The set definition of ray designates a "half-line" in both the Euclidean and Spherical planes (see Figure 1). The definition of ray is *exact* because it describes the spatial object via a precise body of theory: numerical equations and set theory. The definition proves *fruitful* in the subsequent theory developed from the definition. This final criterion assessing the definition based on its ability to yield formal proof relates closely to Antonini and Mariotti's (2008) notion of "theorems" as a triad of statement, proof, and associated body of theory. A definition's quality as an explication is intimately related to the body of theory and proofs built upon that definition.

These definitions are by no means similar in every way to the intuitive concepts they explicate. Distance intuitively corresponds to the space between two points or the distance travelled along a path, but Blau (2008) simply defines distance as a function that obeys certain axioms (positivity "AB≥0", symmetry "AB=BA", definiteness "AB=0 iff A=B"). Rays generally carry a notion of "direction" that is absent from the set-theoretical definition. The second point "B" in a ray is arbitrary, but the definition appears to distinguish it from other points in the ray. The theorem "Given $\text{C}\square\overrightarrow{\text{AB}}$ such that $0<\text{AC}<\omega$, $\overrightarrow{\text{AB}}=\overrightarrow{\text{AC}}$" explicates the arbitrarity of the second point in the ray. The provability of such theorems simultaneously establish the fruitfulness of the explication of ray and restore similarity between formal and less formal meanings (what Weber, 2002, called "proofs that justify axiomatic structure").

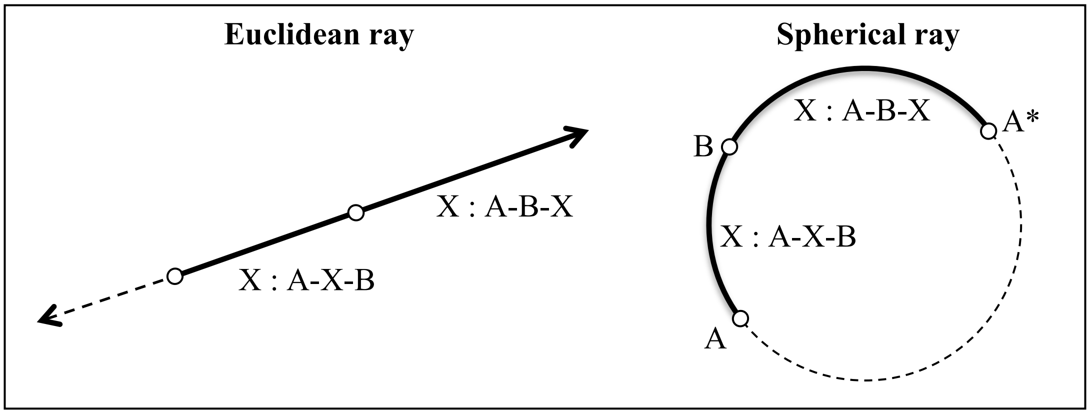


Figure 1. Imagistic representations of rays in the Euclidean and spherical planes.

The episodes in this paper deal with some of the basic proofs that establish the organization of points on a line. Since the geometry is not coordinatized, points are located on lines according to betweenness relations with other points. Two main axioms (Table 2) establish these properties of arrangement: the Betweenness of Points axiom (BP) and the Quadrichotomy axiom (QP). Axiom BP formalizes two basic phenomena on Euclidean and Hyperbolic lines. First, distances between collinear points should be additive: $\mathrm{AB} + \mathrm{BC} = \mathrm{AC}$ (which defines a "betweenness relation"). Second, two points divide the line into three sections such that points in each section satisfy a specific betweenness with the two given points. Because students prove that betweennesses are symmetric (A-B-C ⇔ C-B-A) there are only three distinct betweenness relations among three points, and those three relations fully characterize locations on Euclidean and Hyperbolic lines. Axiom QP captures two other phenomena on lines, including Spherical lines. First, rays with the same endpoint(s) heading in opposite directions cover a line. The betweenness relations in the conclusion of QP match those in the definition of rays $\overrightarrow{\mathrm{BA}}$ and $\overrightarrow{\mathrm{BC}}$ for this reason. Second, given three points on a line (with A-B-C), every other point can be located according to one of four betweenness relations.

Betweenness of Points Axiom	Quadrichotomy Axiom for Points
If A, B, and C are different, collinear points and if $\mathrm{AB} + \mathrm{BC} \leq \omega$, then there exists a betweenness relation among A, B, and C.	If A, B, C, X are distinct, collinear points, and if A-B-C, then at least one of the following holds: X-A-B, A-X-B, B-X-C, or B-C-X.
B-A-C A-B-C A-C-B A C	X-A-B A-X-B B-X-C B-C-X A B C

Table 2. The Betweenness of Points (BP) and Quadrichotomy (QP) Axioms.

An initial proving episode

The pilot study consisted of a series of weekly task-based interviews with a pair of volunteer students from the geometry course. Though I had taught the course previously, I was the interviewer and not the instructor of the course. The instructor was a very experienced teacher who had taught the course at least 20 times and received awards for his instruction. I observed all class meetings and kept extensive written field notes. All interviews were video recorded and audio recorded with a Livescribe© pen. During the second interview, Kirk and Oren sought to prove that the condition $\omega = \infty$ renders QP a theorem rather than an axiom. The instructor assigned the task as a challenge homework problem with the hint that students should use proof by contradiction. I identified four arguments by which the students tried to prove the claim, each of which reveal their evoked meanings for various elements of the body of theory.

First, Kirk and Oren oriented themselves to the task by identifying the claim and assumption towards contradiction. The hypotheses of Axiom QP mention four distinct, collinear points with one betweenness among three of them. The Axiom concludes that at least one of four betweennesses must hold. Kirk paraphrased the assumption toward contradiction saying, “Well it says that, ‘Assume that there's no betweenness relation among any points.’” Kirk jumped from negating four betweenness relations to saying there were none, even though there are 24 possible betweenness relations among four points (12 up to symmetry). Oren questioned Kirk’s interpretation because A-B-C was true by hypothesis. They also noticed the key idea (Raman, 2003) that the sum of any two distances is less than $\omega = \infty$. This meant Axiom BP could be applied to any three distinct, collinear points guaranteeing a betweenness relation. Kirk’s first argument might be summarized as, “The assumption toward contradiction assumes betweenness relations false while BP says they are true leading to a contradiction.” While this is the key idea of the intended proof, Kirk did not attach any inferential meaning (Thompson, 2012) to particular betweennesses, leading him to overgeneralize that none held.

Next, Oren began a syntactic argument analyzing the algebraic consequences of assuming the betweenness A-X-B false. Negating the definition directly yielded $\mathrm{AX} + \mathrm{XB} \neq \mathrm{AB}$. Combining this with the Triangle Inequality, Oren concluded that $\mathrm{AX} + \mathrm{XB} > \mathrm{AB}$. Oren then tried to assimilate this line of reasoning into part of the proof they had recently completed of the triangle inequality. This proof approach did not apply in the current case because it depended upon having finite ω, but it took Kirk and Oren some time to abandon the line of argument.

In the course discussing that second argument, Kirk revealed that he was conflating “there exists a betweenness relation among A, B, and X” with the relation A-B-X, though the phrase entails three possible relations. He also asked whether their completed proof would establish that all four betweennesses in the conclusion of QP would be simultaneously true. The class had already proven the Unique Middle Theorem that states that among three points only one can be in the middle (“If A-B-C, then A-C-B and B-A-C are false”). Also, spatial intuition generally implies only one point can be “in the middle.” Despite this, Kirk was unperturbed in conjecturing that X-A-B, A-X-B, B-X-C, and B-C-X were all going to be true.

Next, based on the interviewer’s clarification of the meaning of “there exists a betweenness among A, B, and X”, Oren developed a third argument comparing the conclusion of BP to their assumption expecting to quickly find a contradiction. He said:

> “Well since they are finite and by axiom BP would imply that there exists a betweenness relationship, relationship such that A-B-X, B-A-X, or A-X-B. But by our [assumption], none of those are true, I think. So [comparing the written lists] X-A-B, A-X-B, there, and A-X, wait, A-B-X. Woah, wait. […] Did I do something weird? A-B-X, oh A-B-X isn't in here, sorry. I thought we had A-B-X was false. Do we know that A-B-X is false?”

While Oren’s discovery did lead in a more productive direction because they could now conclude that A-B-X was true by Axiom BP, it is interesting that up to this point (20 minutes into the proving process) neither Oren nor Kirk compared the particular betweennesses in the conclusion of QP to those implied by Axiom BP. Oren was thus genuinely surprised (and disappointed) when his argument did not lead to a contradiction. Without attending to particular betweenness relations, both students overgeneralized the scope of their initial assumption.

Neither student to this point drew any diagrams in order to interpret the betweenness relations in terms of location or spatial arrangement. To further investigate how Kirk and Oren were interpreting Axiom QP, the interviewer asked them to explain the meaning of the axiom:

K “I think it's any four distinct points, like say A,B,C,X, if you have one betweenness relation, then you're gonna get another one, for a separate three points. I think.”

O “Yeah. There has to be at least one. Does it say at least one, yeah at least one. I think it's either, I don't think three can exist […] I think it has to be one, two, or four. […]”

K “Well no no, I don't think all four of these can hold. […] Because shouldn't there only be four total then? For four points, A,B,C, & X? Like four betweenness between them. […]”

I “But why did you say there can't be just three?”

O “Cause I think if you prove three of them then the fourth one is necessarily true. […]"

K “I know, that's what I'm getting at, so you know that there's gonna be four of these that are going to be true and there's five total. There's this one that we assume's true [A-B-C], then there's four more here [X-A-B, A-X-B, B-X-C, & B-C-X].”

This fourth, combinatorial argument (or conjecture) further confirmed that Kirk and Oren’s meaning for betweenness was incompatible with the spatial intuition of QP dividing the line into four distinct segments. Oren’s claim that three betweennesses implied four likely resulted from an analogy with the class’ Rule of Insertion that said “If A-B-D and B-C-D, then A-B-C-D,” which the instructor had paraphrased “two gives you four.” Kirk now questioned this claim by pointing out that the statement of QP mentioned five betweennesses rather than four. However, their counting strategy never led them to reason semantically about impossible spatial arrangements or syntactically about contradictions to the Unique Middle Theorem. As soon as the interviewer presented a line diagram, the students revised their claims saying only one of the four betweennesses in the conclusion of QP would hold (“It’s here, here, here, or here.”).

Need for explication

To analyze Kirk and Oren’s mathematical activity, I shall introduce what I call the “explication space.” This “space” is a heuristic for analysis and instructional design and should not be taken as a model of student conceptions or thinking per se. As was stated in the introduction, the lens of explication assumes a relationship between less formal and more formal meanings. I apply Carnap’s four criteria for explication as metrics for “measuring conceptual distinctions” between students’ formal and informal meanings within a conceptual domain.

First, as Weber and Alcock (2009) used Goldin’s (1998) theory of representation systems to distinguish syntactic and semantic reasoning, I define students’ more formal and less formal mathematical meanings according to the representation system in which they exist. My use of representation systems is constructivist in the sense of associating inscriptions, meanings and activities according to their relationships in students’ activity. However, because I also must compare students’ mathematical activity to the instructor’s intended learning progression, I rely on the instructor’s meanings (rather than students) to operationally constitute the formal mathematical representation system in the classroom. According to Weber and Alcock (2009), any use of diagrams or images for proving constitutes a shift from the *representation system of mathematical proof* (RSMP) into an alternative, less formal representation system. In such a case, students correspond more formal meanings and expressions with less formal meanings.

This semantic shift “out of the RSMP” occurs with all elements of mathematical theory: definitions, theorems/axioms, and proofs. For definitions, students often shift from the symbolic or verbal notation of the formal definition to an example represented by a generic or particular Diagram. This may be characterized as shifting from the concept definition to part of the concept image (Tall & Vinner, 1981). Regarding theorems, a student may seek to prove a theorem by reasoning about an intuitive paraphrase of the statement claim (Author, year). This ostensibly entails translating the statement from the RSMP into a less formal verbal representation system (much as Kirk paraphrased the assumption toward contradiction as saying “there's no betweenness relation among any points”). I use the term “ostensibly” because previous studies reveal that students often do not recognize that any “translation” occurred (Author, year). From an instructional standpoint, this simply means that the instructor’s intended meaning for the statement and Kirk’s meaning were problematically different. Paraphrasing can also be productive as when Kirk interpreted the statement of QP via the line diagram (Table 2) as saying the point X must be “here, here, here, or here”.

While such “semantic shifting” moves from a formal element (definition or statement) to some less formal correlate, the relationship is reflexive in the sense that formal elements of theory may explicate less formal classes, concepts, patterns, or arguments that are more familiar to the learner. The act of defining often requires identification of a property that designates a class of mathematical objects (Alcock & Simpson, 2002). Creating theorems involves explicating a pattern or mathematical phenomena (Stylianides & Silver, 2009). Axiomatizing requires explicating the necessary conditions for mathematical structures to emerge (Krygowska, 1971). When a student develops an argument or informal proof (Antonini & Mariotti, 2008; Pedemonte, 2007) toward formally proving a claim, they often do so based on less formal meanings or understandings without thorough warrants or details. Arguments are often informal inasmuch as they draw upon less formal meanings for a category, paraphrases of claims, or imagistic representation systems. Weber and Alcock (2009) provide a more detailed set of criteria that distinguish proof in the RSMP from argument, but for the current study the most pertinent aspects of formalizing an argument are a) arguing from concept definitions, b) identifying warrants from the body of theory, and c) adapting to valid logical forms.

Figure 2 visually represents this “space” relating the formal and less formal elements of a body of mathematical theory in their respective representation systems. The left column represents definitions, the middle column represents theorems or axioms, and the right column represents proof. The upper row represents the explications for the lower row’s less formal elements. The upper elements are related to their neighbor below by *similarity*. The upper elements are distinguished from the entry below them by their *exactness*. Finally, the *fruitfulness* of definitions, axioms, and theorems depends upon their ability to yield appropriately explicated elements to their right in the array (definitions must help articulate theorems that must be provable based on those definitions). I also distinguish between whether a formal element entails meanings compatible with its informal counterpart (*similarity*) and whether the informal conception or representation is compatible with its formal counterpart (*co-similarity*).

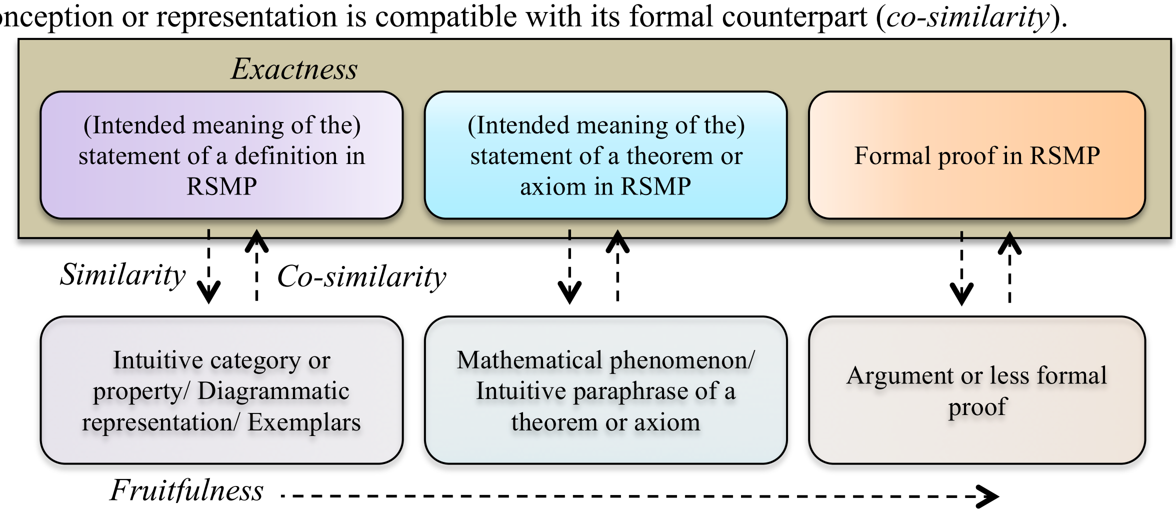


Figure 2. The space of cognitive elements involved in proof-oriented mathematical thinking.

(The need for) explication in the proving episode

As was stated above, the spatial meaning of "betweenness" is the relative location of points while it formally entails equality among three distances. Accordingly, Axiom QP informally means that points (X) on a line with points A, B, and C are in one of four locations. Formally, it guarantees that the distances satisfy one of four equations. Spatially, only one of the four betweennesses in the conclusion of QP should hold (unless X is the antipode of B). According to the Unique Middle Theorem in the class body of theory, at most two are possible. Thus, there are at least two (somewhat distinct) sets of meanings for betweenness that could be used for semantic or syntactic reasoning about the given claim. However, neither Kirk nor Oren referred to spatial meaning nor were they able to fruitfully use the equality meaning and class theorems. Ultimately each of their four arguments entailed flaws that inhibited progress toward valid proof.

I conjecture that Kirk did not pay attention to the particular betweennesses because they entail heavy cognitive load as mere strings of letters (without some meaning for drawing inference). As such, he paraphrased the assumption toward contradiction (no betweennesses hold) and the consequences of the condition $\omega = \infty$ with Axiom BP (betweennesses must hold), which suggested a pathway toward a contradiction. This key idea (Raman, 2003) is actually very important for producing the intended proof of this theorem, and is suggested by the textbook's presentation of the task. However, as Kirk rephrased these formal statements, he significantly changed the intended meanings leaving him unable to construct proof. Stated in the language of explication, Kirk's paraphrase was not *co-similar* to the intended meanings of the formal statements about which he reasoned.

Oren's syntactic argument could be considered an example of using a proof frame (Selden & Selden, 1995) or adapting a previous proof technique (Weber, 2002). This strategy is often useful and has been endorsed in proof literature, but in the present case Oren's reasoning was frustrated by lack of logical *exactness* when he tried to assimilate the new statement into the former proof frame. Without any guiding semantic interpretation of betweennesses, Kirk and Oren were very slow to recognize the *fruitlessness* (and logical fallacies) of their arguments. In this case, Oren's reasoning about the definition of betweenness was *exact* but not *similar* to the spatial meaning of between. Furthermore, this lack of inferential meaning led Kirk to conjecture that they would simultaneously prove X-A-B and A-X-B, which is neither spatially nor formally appropriate.

The third argument led Oren to attend to particular betweenness relations among three given points. However, without directly comparing the conclusion of QP and the entailments of "there exists a betweenness among A, X, and B", Oren fully expected a contradiction. This further emphasized that he lacked an effective semantic meaning for the conclusion of QP. This made him rely on purely syntactic operations directly comparing the strings of letters. While this is not problematic per se, he attempted to form a new conjecture (their fourth argument) based on simply counting betweennesses. Again, this combinatorial approach is not unproductive or unfruitful a priori. The students were likely motivated by their instructor's paraphrasing action ("two gives you four"). This line of reasoning is limited as it only assigns truth-values to betweennesses and entails no inferential meaning (distance equality or spatial location) by which to assess the verity of the claims being made.

Once the interviewer presented Kirk and Oren with a diagram, they quickly assigned spatial meaning to the conclusion of QP. Prior to that, Kirk and Oren displayed very fragmented meanings for betweenness rendering much of their reasoning *fruitless*. Their understanding of formal statements such as QP was not *similar* to the motivating intuition drawn from the diagram in Table 2. One might even ask in what manner any of Kirk or Oren's activity represented "geometric reasoning." Their paraphrases of the formal statement were not *co-similar* to the intended meanings. Also, because they used incompatible proof frames, the logic of Kirk and Oren's reasoning tended to lack *exactness* in the sense of validity.

To avoid possible criticisms of the chosen proving episode, it is important to make three points about Kirk and Oren's learning. First, Kirk and Oren received good geometric instruction from a very experienced and capable instructor. However, treating the formal elements of theory as explications of intuitive geometric concepts was not necessarily a guiding goal for the instructor. The instruction was more compatible with traditional proof-oriented instruction (Weber, 2004), which often gives logical organization primacy. Second, Kirk and Oren's fragmentation of meaning was not isolated to this proving episode or to the absence of diagrams. In later episodes in which they reasoned about diagrams, Kirk often reasoned about betweenness as order ("A then B then C") introducing directionality and again losing co-similarity with formal definitions. Oren often metonymized geometric objects struggling to distinguish distances (numbers) from segments (sets) from betweennesses (properties), violating exactness. Third, Kirk and Oren passed their geometry course indicating successful learning overall.

Explicating planar geometry through reinvention

These interviews motivated my use of a reinvention/explication lens to help students coordinate formal and spatial meanings in axiomatic geometry. As was mentioned previously, proving in this setting seems to necessitate coordinating semantic and syntactic reasoning. I drew my primary design heuristics for reorganizing the course from the RME tradition, including *guided reinvention* and *emergent models* (Freudenthal, 1973; Gravemeijer, 1994). I first piloted reinvention activities first in a full semester of the axiomatic geometry course for which I was the instructor. To further refine and document the influence of the activities, I then conducted a quasi-teaching experiment in conjunction with a geometry course taught by Kirk and Oren's instructor. My participants continued to receive geometry instruction during their class meetings, such that I did not control that portion of their learning. To account for this, I guided the group of 3 volunteers in reinventing definitions and axioms in advance of their appearance in the lectures. In the following sections, I outline Gravemeijer's framework for guided reinvention and descibe how I adapted that frame for axiomatic geometry. Using data from the teaching experiments, I provide an example of one such reinvention progression for the concept of "line". I then present an account of the teaching experiment participants proving "If $\omega = \infty$, then QP" after they had reinvented most of the constituent definitions and axioms.

Gravemeijer's (1994) stages of mathematical activity

The guiding philosophy of RME descends from Hans Freudenthal's (1973) emphasis on mathematics as a human activity. He strongly advocated that students' mathematical learning develop from their mathematical activity. He criticized imposing the *products* of experts' mathematical thinking on students instead of engaging them in the organizing *activities* from which those products came. For Freudenthal, mathematical structure comes not from the abstractness of mathematical language or formal logic, but from students' structuring their mathematical activity. To express this shift from product to activity, Freudenthal was fond of verbing nouns: mathematics became mathematizing, definitions – defining, axioms – axiomatizing, theorems – theorizing, proofs – proving, etc. One of the prime heuristics for making these processes actionable in the classroom is *guided reinvention*. Through reinvention of key mathematical concepts, students gain access to and training in the processes of mathematical activity as well as understanding of the products of that activity. Successful reinvention is an instance of psychological explication whenever students already possess some intuitive or less formal conception of a mathematical structure as is the case in planar geometry. Reinventing an axiomatic theory of planes entails creating *similar*, *exact*, and *fruitful* definitions and axioms from which precise theorems can be formulated and proven.

Instructors guide reinvention by supporting students in three basic processes: 1) students engage in experientially real activity, 2) they reflect on that activity to apprehend structure, and 3) they adapt that structure to further experiences. The latter two components came to be known as students' creating a *model-of* and *model-for* their mathematical activity, respectively. Gravemeijer (1994) identifies four stages in which these models develop. Students' initial exploration of an experientially real situation constitutes the *situational stage*. This may involve play, a context problem, or even students' prior learning. As students structure their activity through reflection, they construct a model-of their activity in the situation. This is the *referential stage* in which students reason about concepts and inscriptions in ways that continue to refer to the original situation. Furthermore they begin to structure their thinking about the situation in terms of the model. As students continue to use and elaborate the model, their activity may cease to depend upon the original task setting at which point they pass into a *general stage* of activity. This transition often accompanies the students' assimilation of a distinct situation into the model or exploration of questions that do not reflect the conditions of the original situation. At that point, the student's *model-of* the situation has become a *model-for* further mathematical activity and has become an independent experiential entity. Gravemeijer points out that transition to this stage constitutes the students' creation of a *new mathematical reality*. Finally, once students reason within this new mathematical reality without recourse to the first situation or the activities drawn from it, they have passed into the *formal stage* of mathematical activity.

Reinventing axiomatic geometry while riding the subway

In the teaching experiment, the students went through multiple iterations of Gravemeijer's stages of reinvention/explication with respect to different components of the body of theory. In the course of explicating notions such as line, distance, and between, students also developed an overarching model of the concept of geometric plane, which could accommodate a variety of example planes. This was a primary goal of the teaching experiment because I observed during pilot studies that students did not see anything geometric about finite geometries as represented by sets of letters and tables of distances. Also, students tended to reason very differently about spatial diagrams versus analytical inscriptions. To provide an experientially real situation for mathematizing finite geometry, I asked students to consider a subway system with seven stations. Already being familiar with basic elements of Euclidean, Spherical, and Hyperbolic geometries from their class meetings, participants were asked to identify points, lines, and distances in this "geometry". The teaching experiment group agreed that stations should be points and that subway lines would be lines. Most students wanted the distance between points to be "number of stops." This situation supported the students' initial defining activity.

Motivated by the explication criteria of *exactness* in mathematical definitions, I posed the question "What kind of mathematical object is a line?" along with a set of possible answers. The group suggested lines could be equations, graphs, functions, or sets. The final option emerged last and reluctantly. This was situational reasoning because most students' answers reflected their concept image of line developed through years of algebra and Euclidean geometry. The students now had four possible explications of "line." When asked which best fit with lines in Spherical, Hyperbolic, or Subway geometries, students agreed that the notion of set most easily accommodated each kind of line (referential activity). The notion of "set of points" is *similar* to students' notion of line in the sense that it can describe familiar lines. However, it is not *co-similar* in the sense that students hold many meanings of line that are not entailed in or even related to this definition (straightness, extension, density, etc.). To distinguish this new definition from intuition, I asked the group to create definitions for the terms parallel and intersecting. Students first suggested that lines intersect if they "cross", but admitted they could not explain what this meant in terms of sets. The students then suggested that two lines (sets) intersected if a point was in both sets and were parallel if they shared no common points (without guidance they did not connect this to the set operation "intersection"). This defining activity guided students to differentiate their broader concept image of line from the formal definition (set of points) and its entailments. This differentiation began to move them toward more general activity.

Formalizing mathematical activity leading to proof

Based on these initial explorations and thinking about the subway situation, the participants created basic "rules" for lines and distances that constituted the initial axioms of planar geometry. These included that distances are positive, distances are symmetric, distances are additive (AB+BC=AC), the triangle inequality, every two points must lie in a line, and every line must have at least two points. Globally, the emergent structure comprised of these definitions and rules began to form the students' axiomatic model-of geometric planes.

Later, after students were comfortable creating new finite geometries using letters for points, sets of letters for lines, and a table for the distance function, the were charged with creating a finite geometry that violated the triangle inequality. This was no trouble for the students in the analytic representation system of sets and tables. They defined three lines each containing two points and assigned distances: AB=1, BC=1, and AC=4. When asked to produce a diagram for this plane, Benjamin drew an open triangle keeping all of the line segments straight (Figure 3). While this diagram represented distance appropriately, two different endpoints corresponded to the point B. Matthias instead drew a "waffle cone" that connected all of the appropriate segments, but curved the line AC to represent the greater distance. Benjamin protested that Matthias' line broke the vertical line test.

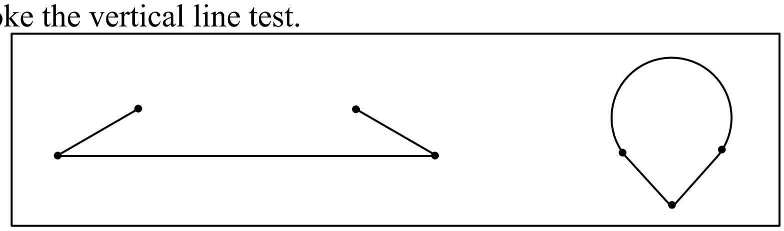


Figure 3: Benjamin and Matthias' diagrams of planes violating the Triangle Inequality.

This episode reflects several things about the students' set explication of line. The group operated in the *general stage* with regard to their notion of line in the representation system native to finite geometries. They were unperturbed in defining lines with only two or three points in set notation. This constitutes a distinct concept of "line" from what they had prior to the course, indicative of a new mathematical reality. However, when instantiated spatially Benjamin reverted to *referential reasoning* in which lines must be straight like Euclidean lines. This suggests some "context dependence" to the formality of students' reasoning as they reinvent/explicate intuitive concepts. Because Benjamin endorsed the triangle inequality as an axiom, he likely maintained straightness at least partially to emphasize the counterintuitiveness of violating the inequality. Matthias accommodated his spatial meaning of line to allow curvature indicating general activity independent of the Euclidean origins his conception of line. Benjamin's protest according to the "vertical line test" clearly refers back to his situational concept image developed in prior instruction. Matthias showed that he was able to adapt his imagistic reasoning to establish *co-similarity* with the strange types of lines afforded by the axiomatic body of theory, unlike Benjamin's imagistic reasoning that lacked *co-similarity*. Because the class' body of theory provided no warrant for the straightness of lines, Benjamin drew upon his prior learning to argue against Matthias' "Waffle Cone Plane" (as they called it).

After a few weeks of further mathematical activity, I asked the participants to prove the claim, "If A-B-C, then $\overleftrightarrow{AB} = \overrightarrow{BA} \cup \overrightarrow{BC}$." When viewed spatially as in Table 2, students struggle to understand what must be proven because the referents to both sides of this equation are the same part of the diagram, which strongly stymied Kirk and Oren's progress on this same task during the pilot study. Unlike Kirk and Oren, the participants in the latter study were able to formulate a valid proof. An important step in their progress was formulating an assumption toward contradiction that there existed a point X in the line that was not in the union of the rays (they separately proved that $\overleftrightarrow{AB} \supseteq \overrightarrow{BA} \cup \overrightarrow{BC}$). This step depended on interpreting the conclusion of the claim as equality between sets. At this point, I claim the study participants were acting in the formal stage with regard to their conception of line because they could reason about lines as sets without perturbation. As a result, it is valid to say they had psychologically explicated their understanding of line. Their conception of line was *similar* to Euclidean lines in the sense of being able to assimilate them into the set meaning. Matthias could also use diagrams as heuristics for reasoning about lines in a manner that maintained *co-similarity*. Their reasoning was *exact* in the sense of appropriately maintaining the structure of set-theoretic relationships. Also, their reasoning became *fruitful* when they produced valid proofs of geometric claims.

Proving "If $\omega = \infty$, then QP" – identifying a proof frame

The overall hypothesis of the teaching experiment was that students who see formal definitions/axioms as explications of intuitive or less formal geometric notions will be better able to coordinate semantic insights and heuristics with syntactic proving activities in the exact body of theory developed in the class. The students in the teaching experiment (Group 2) were thus asked to prove some of the same claims that had been posed to Kirk and Oren (Group 1) to trace how they drew upon their explication activity for proof production. For comparative analysis, I shall describe Group 2's reasoning about "If $\omega = \infty$, then QP".

When first presented with the claim, Benjamin suggested that the group write out the exact statement of QP, which they did using the textbook. Having "reinvented" QP from their analysis of Euclidean and Spherical lines, Matthias also wrote down how the betweennesses in the conclusions of QP related to membership in the rays $\overrightarrow{BA}$ & $\overrightarrow{BC}$. Matthias introduced an initial argument suggesting that they write the definitions of the two rays and try to prove that the point X must be in one of them. Because they had previously assumed QP as an axiom, the group struggled to identify the hypotheses and conclusions for the claim to be proven. Benjamin also asked why it was important to assume $\omega = \infty$. Matthias tried to clarify saying, "So really what we have to prove is that X is in the union of these two [rays] because that would require one of these four [betweenness relations]. Cause if X is outside that union, that is the only time it's not one of these four, so basically it has to always be in there." Benjamin agreed and pointed out that Matthias' explanation began a proof by contradiction. In this way, the group identified a sufficient condition for proof (X is in the union of rays) and a proof technique (by contradiction).

Next, Matthias began performing syntactic operations on the assumption toward contradiction. He translated $X \notin \overrightarrow{BA} \cup \overrightarrow{BC}$ into $X \notin \overrightarrow{BA}$ and $X \notin \overrightarrow{BC}$. Benjamin struggled to understand how X could not be a member of either ray claiming it contradicted the definition of ray. Matthias however shifted the focus of the conversation by introducing the same key idea that Group 1 used. Because the sum of any two distances is finite and $\omega = \infty$, Axiom BP guaranteed the existence of a betweenness relation among any three points. He explained this point several times before his group mates assented to the argument. Then, Matthias applied this to points A, B, and X because the conclusion of QP included two such betweenness relations. He explained that since they had assumed two of them false, they could conclude A-B-X (and X-B-C).

Proving "If $\omega = \infty$, then QP" – rooting out spatial impossibilities

Prior to this, no line diagrams appeared on the board. Benjamin introduced a spatial (semantic) interpretation of the given claims. Figure 4 presents the diagram he drew as he said,

> "Well, A-B-X is contained in the ray $\overrightarrow{BC}$. I mean, intuitively, yeah. At least, Euclidean. If you draw a picture of a Euclidean line. Cause I'm just looking at it, if you've got [drawing Figure 4]. So A-B-X, your X is here, or your X is to the right of B. It could be anywhere [to the right]. And for C-B-X, it goes C-B [moving his hand right to left] your X is here. Somewhere along here [running hand over left half of diagram]. It could be past A. […] But I don't know, so let's look at it on the sphere. [Matthias points out that $\omega = \infty$ excludes the sphere.] Well how do we prove it without using, we can't just look at it on a Euclidean line. We have to look at it on all of our planes. I mean our end goal is to prove that A-B-X is contained in the ray $\overrightarrow{BC}$ and C-B-X is an element of […] the ray $\overrightarrow{BA}$. That's the end goal for the contradiction. How do we get there?"

The group now attempted to apply the Rule of Insertion to the betweenness relations they knew (A-B-C, A-B-X, and X-B-C), but no two of these satisfy the hypotheses of that theorem.

At this point, the group became stymied. They recognized the counter-intuitiveness of their current claims, but did not know how to achieve a contradiction. They resorted to searching the text for tools. The next step in the intended proof requires applying BP to the points A, C, and X for which they did not currently have a betweenness relation. To lead the students to this, I asked them how many betweenness relations were possible among four points (A, B, C, and X). I also asked this to compare their reasoning to Kirk and Oren's combinatorial reasoning. Benjamin quickly suggested either 12 or 24 explaining that it should be "four choose three." Unclear whether there were 12 or 24, the group produced an exhaustive list organized by first point (6 possible) and then second points (2 possible). After some interchange, they identified that the only betweenness relations that were not yet determined in their proof of QP involved A, C, and X. I suggested that they should examine the consequences of each betweenness among these points. Last, the group successfully employed their previous proof strategy by applying the Rule of Insertion to the newly assumed betweenness relations and showing that each entailed a contradiction to the Unique Middle Theorem. Once they proved each possible betweenness among A, C, and X entailed a contradiction, they completed the proof.

Figure 4: Benjamin's line diagram for establishing a contradiction.

The effects of explication on students' proof activity

Group 2 showed a marked difference in their proof production as compared to Group 1. Group 2 began by orienting themselves semantically to the claim recognizing that the two betweennesses in QP matched the definition of two opposite rays whose union should cover the line. This led Matthias to propose a sufficient condition for proving QP: $X \in \overrightarrow{BA} \cup \overrightarrow{BC}$. Because this claim appeared intuitively obvious, the group agreed they should use proof by contradiction. Unlike Group 1, Group 2 quickly interpreted the betweenness relations in QP as statements of location or spatial arrangement allowing them to gain insight into the verity of the claim.

Both groups identified the key idea that $\omega = \infty$ allows any three collinear points to satisfy the hypotheses of Axiom BP. Group 1 was limited in their use of this because their lack of meaning for betweenness relations non-trivially changed the conclusion of BP from "there exists a betweenness relation" to "all betweenness relations will hold" [my paraphrase]. On the other hand, Matthias quickly noted that of the three betweenness relations possible among A, B, and X, two were assumed false. When Group 2 explicated their intuition of "between", Matthias developed a rich concept image of betweenness relations that allowed him to interpret and coordinate the strings of letters in ways that Group 1 lacked. This was confirmed by Group 2's very productive exploration of the combinatorics of betweenness relations among four points.

Benjamin's reasoning in the QP episode is noteworthy in several ways. First, he repeatedly showed how he used Euclidean line diagrams to interpret the written conditions. This means Benjamin consistently showed preference for semantic reasoning. Previous studies (Weber & Alcock, 2009; Alcock & Simpson, 2004) suggest that semantic reasoners often reap intuitive insight and conviction from their reasoning, but may lack motivation to formalize arguments. In the language of emergent models, the "waffle cone" episode indicated that in imagistic representation systems Benjamin used referential reasoning rooted in the Euclidean situation. Benjamin's activity in proving QP suggests several caveats to these classifications of his reasoning. Benjamin's semantic interpretation led him to the insight that A-B-X should place $X \in \overrightarrow{BC}$, which was not obvious to Matthias who operated more syntactically. Productively, Benjamin qualified this insight in several ways. He knew that the diagram did not constitute proof, saying "intuitively" and "at least Euclidean". He then suggested they test the claim in a different geometric space. He finally formulated his semantic insight into a new subconjecture sufficient to reach contradiction: "[the point X in] A-B-X is contained in the ray $\overrightarrow{BC}$". In this way, Benjamin qualified his semantic insights acknowledging the need for proof.

Also, Benajamin accommodated his line diagram to the constraints of the task by drawing auxiliary lines in a different color that represented the two rays. As was previously mentioned, in a straight line diagram the same sections of line represent the rays $\overrightarrow{BA}$ and $\overrightarrow{BC}$ and their carrier line $\overleftrightarrow{AB}$. He altered his diagram so that the rays and line had different but related representations, indicating his active differentiation of the two sets: $\overleftrightarrow{AB}$ and $(\overrightarrow{BA} \cup \overrightarrow{BC})$. He articulated part of his claim using the phrase "an element of" indicating set structure rather than the common "X is on the ray" indicating spatial location. This suggests Benjamin's had transitioned into formal geometric activity because he reorganized his reasoning about the Euclidean plane to reflect the structure of the class' geometric body of theory. I claim Benjamin psychologically explicated his intuition of line because he formalized his reasoning rendering it ready for proof production. In addition, the cognitive history of his reinvention activity allowed him to maintain the intuitive insights reaped from spatial reasoning in the Euclidean plane. As such, even Benjamin's analytic activity maintained a geometric nature as it expressed statements about spatial arrangement.

Concluding Remarks

While it remains unclear to what extent Kirk and Oren's mathematical activity should be considered an explication of geometric reasoning, my ongoing investigations of reinvention affirm Sjogren's (2010) claim that formal proof can be viewed as the explication of less formal proof. However, this is primarily true after students have first (a) explicated intuitive meanings through defining and (b) explicated mathematical phenomena through axiomatizing and theoremizing. As Antonini and Mariotti (2008) argue, proof is contextualized within a body of theory and the relationship between informal and formal proof cannot be separated from this larger theoretical milieu. In the context of the reinvention of a body of mathematical theory, explication provides useful criteria for analyzing student difficulties and for developing alternative instructional activities in conjunction with the RME instructional heuristics of guided reinvention and emergent models. While the RME heuristics provide a rich frame, *similarity*, *co-similarity, exactness*, and *fruitfulness* provide more specific language for analyzing students' reasoning and comparing it to the formal meanings instructors intend for them to develop.

Comparing the pilot study to the teaching experiment reveals a marked difference in students' proof production as a result of explicating less formal mathematical meanings. The two groups received comparable mathematical instruction from the same professor, but Group 2 displayed much more coherence in their meanings for key geometric concepts and greater ability to productively draw upon both semantic and syntactic types of reasoning. It is precisely this coordination of intuitive and precise mathematical meanings that I suggest students (a) should benefit from and (b) can develop through appropriate reinvention/explication-oriented activities.

References

Author, A. (year).

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

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.

Antonini, S. & Mariotti, M.A. (2008). Indirect proof: what is specific to this way of proving? *ZDM Mathematics Education, 40*, 401-412.

Blau, H. (2008). *Foundations of Planar Geometry*. Island Park, NY: Whittier Publications.

Carnap, R. (1950). *Logical foundations of probability*. Chicago, Ill: University of Chicago Press.

De Villiers, M. (1998). To teach definitions in geometry or to teach to define? In Olivier, A. & Newstead, K., *Proceedings of the 22nd Annual Meeting of the International Group for the Psychology of Mathematics Education.* (2, pp. 248-255), Stellenbosch, South Africa: University of Stellenbosch.

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht: D. Reidel Publishing.

Goldin, G. (1998). Representational systems, learning, and problem solving in mathematics. *Journal of Mathematical Behavior, 17*, 137–165.

Gravemeijer, K. (1994). *Developing Realistic Mathematics Education*. Utrecht: CD-β Press.

Krygowska, A.Z. (1971). Treatment of the Axiomatic Method in Class. In Servais, W. & Varga, T. (Eds.) *Teaching School Mathematics.* (pp. 124-150), London: Penguin-Unesco.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics, 66*(1), 23–41.

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.

Selden, J. & Selden, A. (1995). Unpacking the Logic of Mathematical Statements. *Educational Studies in Mathematics, 29*(2), 123-151.

Sjogren, J. (2010). A Note on the Relation Between Formal and Informal Proof. *Acta Analytica, 25*, 447-458.

Stylianides, G. & Silver, E. (2009). Reasoning-and-Proving in School Mathematics: The Case of Pattern Identification. In M. Blanton, D. Stylinaiou, & E. Knuth (Eds.), *The teaching and*

learning of proof across the grades: A K-16 perspective. (pp. 235-249). New York: Routledge.
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.
Thompson, P. (2013). In the absence of meaning… In Leatham, K. (Ed.), *Vital directions for research in mathematics education*. New York: Springer
Weber, K. (2002). Beyond proving and explaining: proofs that justify the use of definitions and axiomatic structures and proofs that illustrate technique. *For the Learning of Mathematics, 22*(3), 14-17.
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. & Alcock, L. (2009). Semantic and syntactic reasoning and proving in advanced mathematics classrooms. In M. Blanton, D. Stylinaiou, & E. Knuth (Eds.), *The teaching and learning of proof across the grades: A K-16 perspective*. (pp. 323-338). New York: Routledge.
Zandieh, M. & Rasmussen, C. (2010) Defining as a mathematical activity: A framework for characterizing progress from informal to more formal ways of reasoning. *Journal of Mathematical Behavior, 29*(2), 57-75.

FORMATIVE ASSESSMENT AND STUDENTS' ZONE OF PROXIMAL DEVELOPMENT IN INTRODUCTORY CALCULUS

Rebecca-Anne Dibbs, University of Northern Colorado
Michael C. Oehrtman, University of Northern Colorado

One of the challenges of teaching introductory calculus is the large variance in student backgrounds. Formative assessment can be used to target which students need help, but little is known about why formative assessment is effective with adult learners. The purpose of this qualitative study was to investigate which functions of formative assessment as described by Black & William's 2009 framework. This paper examines case studies of two students: Leonard and Sandra. Although the two students earned similar grades, their varying levels of participation in the formative assessments led to very different conceptual development paths in their introductory calculus course.

Key words: approximation framework, formative assessment, self-monitoring, Zone of Proximal Development

Formative assessments, low stakes assignments given to assess students' current level of understanding, increase student achievement (Black & Wiliam, 2009; Clark, 2011), but little is known about how implementing formative assessments facilitates this achievement gain. The purpose of this research was to study the impact of formative assessment on students' engagement in their Zone of Proximal Development (ZPD) in a calculus course designed with Oehrtman's (2008) approximation framework. Our central research question is: How does formative assessment affect students' engagement in their ZPD and conception of the limit structures as developed in Oehrtman's (2008) approximation framework for calculus instruction?

Understanding how the use of formative assessment affects college students' engagement in their ZPD and development of a particular conceptual structure can advance the theory of formative assessment, which has been most prominently influenced by research in European primary and secondary schools (Black & Wiliam, 1998; 2009). Black & Wiliam's (2009) framework of formative assessments suggests that there are five major benefits of formative assessment (Figure 1).

	Where the learner is going	Where the learner is right now	How to get there
Teacher	**1** Clarifying learning intentions and criteria for success	**2** Engineering effective class-room discussions and other learning tasks that elicit evidence of student understanding	**3** Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	**4** Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	**5** Activating students as the owners of their own learning	

Figure 1. The five purposes of formative assessment (Black & Wiliam, 2009)

We used social constructivism as the theoretical framework for this study, specifically Vygotsky's Zone of Proximal Development (ZPD). There are three characterizations of the ZPD (Vygotsky, 1987). The first is the collaborative ZPD, where students can solve problems in a group that they cannot solve on their own, and the second characterization of the ZPD is that students are within their ZPD if they can solve a problem with scaffolding that they cannot solve on their own (Lave & Wenger, 1991). This report will focus on the third characterization of the ZPD: interplay between students' spontaneous and scientific concepts and scaffolding that supports students in deepening their conceptual understanding. Spontaneous concepts are those that are largely intuitive and loosely structured for students; while students may be able to complete tasks when they have a spontaneous conception, they cannot explain their reasoning (Vygotsky, 1987). Scientific concepts are ones that students learn through formal instruction; the definition and procedures are learned before the intuition develops (Vygotsky, 1987). A concept is rarely purely one conception or the other; students are said to have appropriated a concept when they have possess the both the scientific and spontaneous conception that is accepted by the community of practice (Vygotsky, 1987).

Students are in their ZPD if they are using spontaneous reasoning to understand a scientific concept or vice versa (Vygotsky, 1987). Generally, the ZPD can be identified by determining what students can do but only with assistance. The learner is a peripheral participant in this assessment and subsequent scaffolding, because they are being assisted by a more central member of the learning community (Lave & Wenger, 1991; Smagorinsky, 1995). As the learner gains expertise, scaffolding may be reduced and the learner becomes a more central participant in the community of practice.

We argue that formative assessment enabled instructors to better assess and target the areas in which critical scaffolding was needed and that this process increased both students' self-monitoring of their understanding and opportunities for peripheral participation in the classroom.

Methods

We recruited participants from three sections of introductory calculus utilizing Oehrtman's (2008) approximation framework as a coherent approach to instruction. This framework is built upon developing systematic reasoning about conceptually accessible approximations and error analyses but mirroring the rigorous structure of formal limit definitions and arguments (Oehrtman, 2008, 2009). This paper focused two students' conceptual development on the three multi-week labs developing the most central topics in the course: Lab 3 (limits), Lab 4 (derivatives), and Lab 7 (definite integrals).

The two participants in this study, Leonard and Sandra, are similar at first glance. Both students were physical science majors. Neither student had taken a calculus course before, nor had either enrolled in any mathematics courses in two years. Leonard and Sandra were opposed to working in groups throughout the semester. Leonard and Sandra earned test and webwork grades within two percent of each other on all assignments. They also earned almost the same grade in the course: Leonard earned a 68% and Sandra earned a 72%. The difference in their final grades can be attributed to one factor: Leonard did exactly two post-labs all semester (Pre-lab 3 and Post-lab 4B), while Sandra did every post-lab.

Figure 2 provides a portion of a typical formative assessment. These assignments were given prior to each lab (pre-lab) and after each day of lab work (post-lab). The first questions of our formative assessments were conceptual questions about important aspects of the approximation structures in the current lab (not shown in Figure 2). Two open-ended questions always appeared as the last two questions of every formative assessment. We recorded which concepts each student (n = 46) explicitly stated they did or did not understand and what, if any errors they made on computational questions on each formative assessment.

A summary of students' responses to the questions were used to plan a brief intervention in the next class addressing the problematic issues.

Post-Lab 3A: Locate the Hole (Limits)
Directions: Answer the following questions to the best of your ability. Responses need not be lengthy, but should answer all parts of the question.

1. Which question is your group working on?
2. What have you figured out about the answer so far?
3. Write a short paragraph that answers the following two questions. What mathematical concepts or phrases used so far this week do you recognize from calculus? From other mathematics courses?
4. What questions do you have about the material we have covered so far in class?

Figure 2. A typical formative assessment

This qualitative study centered on a document analysis (Patton, 2002). Our primary sources of data were student documents: formative assessments, homework assignments, and exams of all students in two sections of introductory calculus, with particular attention paid to ten students who participated in at least one interview. During each of the two interviews, one after Lab 4 and the other after Lab 7, participants were presented with copies of their lab write-ups and formative assessments. For each cell in the approximation framework (Table 2), students were asked to explain why they wrote their solution, what help they got during and outside of class, and if there was anything about their solution they would change. The student interviews were triangulated with informal interviews with the instructor and class observation notes. Dibbs observed the classrooms every day pertaining to either the formative assessments of the labs and debriefed the instructors on a weekly basis to obtain their observations of student and classroom learning trajectories (Table 1).

Table 1
Lab 3 Timeline

Day	Description
Monday	Pre-lab 3 given to students
Tuesday	Lab 3 in class (groups); post-lab 3A in evening
Wednesday	Instructor intervention based on post-labs
Tuesday	Lab 3 Jigsaw (groups); post-lab 3B in evening
Wednesday	Instructor intervention based on post-labs
Friday	Lab 3 individual write-up do

We analyzed the data using a set of codes drawn from Vygotsky (1987) that we operationalized in terms of the data after an initial round of exploratory coding (Table 2). We then used the approximation framework code grid (Figure 3) and the code dictionary to classify each cell of the code grid as one or two primary codes from the dictionary for each student's lab write-up. A code was considered appropriate for the cell when the student's explicit statement in the interview concurred with the written work and the observation notes. In the case of conflicting data, the chapter exams were coded as a tiebreaker. The final exam problem based on Lab 7 was also coded for all students as a check on the Lab 7 codes.

Table 2
Coding Dictionary

Code	Type	Abbreviation	Standard of Evidence
Situation-bound reasoning	Spont.	SBR	Student provided correct solution in one context but was unable to provide the correct solution in a new context.
Classification preceded explanation	Spont.	CPE	Student was able to classify parts of the framework correctly (over- and underestimates; error/ error bound) but could not provide an explanation for their thinking.
Increased Quality	ZPD	INQ	Student presented a more complete solution than on the previous approximation lab.
Less Scaffolding	ZPD	LS	Student needed less scaffolding to complete the solution than on the previous approximation lab.
Appropriated	ZPD	APP	Student was able to produce the complete solution in a cell where they were volitional in a previous lab.
Volitional	Sci.	VOL	Student was able to produce the complete solution with no scaffolding, or student answered lab questions in an interview setting.
Plan is right, work is not	Sci.	PRW	Student had the correct strategy (in the context) to produce a solution, but their solution contained at least two mathematical errors.
Learned through instruction	Sci.	LTI	Student was explicitly taught the solution strategy during the lab, office hours, or the intervention.
Unjustified heuristic	Sci.	UH	Student selected a previously successful strategy for solving a problem, but failed to translate to the new context. (eg. Approximating derivatives with y values instead of slopes).
Ventriliquation	Sci.	VENT	Student solution is a paraphrase of scaffolding given to the student in the lab, office hours, or intervention. Student can provide no explanation for the solution.

The post-labs were coded for three things: (1) mathematical errors students made on any calculational questions, (2) noting if the students identified the problems they had with calculations or parts of the lab accurately, and (3) coding all questions by the concept students found troubling. During the intervention, the first author observed the class using three minutes to count student behaviors (paying attention to the instructor, taking notes, texting or other off task behavior) and then spent three minutes recording impression. Those observation notes were coded for changes in participation patterns.

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

Figure 3. Approximation framework code grid

Students' responses to the formative assessments were triangulated with field notes of the classes immediately before and after the lab, as well as their submitted lab reports. Each student's work was coded for particular areas of improvement after the intervention. After the coding was complete, it became clear that despite their similarities, Leonard and Sandra had very different conceptual development trajectories on the labs.

Findings

Leonard worked in a permanent group throughout the semester: Emily who earned an A and Lisa who failed the course. Sandra changed groups each lab, never really finding a group she liked to work with. In this section, I discuss the learning trajectories of each student throughout the semester, starting with Leonard.

Leonard

Leonard is a sophomore chemistry education major. Introductory calculus was the only required mathematics course in his major. Leonard took pre-calculus in high school, but this was the first course where he had seen calculus topics. It had been a year since his last math class, and Leonard admitted that he forgot most of the math covered in that class, especially trigonometry. Leonard did have a well-developed number sense, which he attributed to his science courses, which let him evaluate the reasonableness of his answers. Leonard made far fewer calculation errors than one would expect from a student that did not pass the course, but his weak algebra and trigonometry skills caused Leonard problems throughout the semester.

Lab 3

In the first approximation lab, Leonard's lab was approximately half situationally meaningful reasoning, and half learned through instruction. He had eight cells coded as situation-bound reasoning and eight cells coded as learned through instruction. One cell of the approximation framework was coded as ventriliquation, because Leonard stated in his interview that his response to the final question on the lab was primarily dictation from the lab facilitators, and that he didn't really understand what he was writing. Hence, slightly over half of Leonard's first approximation lab was scientific reasoning and the other half was spontaneous reasoning because of Leonard's familiarity with functions.

The first concept of the approximation framework, the unknown value, is given to students to complete as their pre-lab assignment. Leonard's initial graph was created using an unjustified heuristic that a linear factor in the denominator of a function meant the presence of an asymptote. When looking over the pre-lab in his interview, Leonard stated, "This was the easy one…It was a missing y value in a function, and I've done functions in other classes before." Since the context made sense to Leonard, he was able to make sense of the unknown value parts of the approximation in terms of the context; however, since Leonard was unable to complete this portion of the approximation framework in the next lab, this was situation-bound reasoning.

Table 3.
Leonard's Lab 3 Codes

	Contextual	Graphical	Algebraic	Numerical
Unknown Value	SBR	SBR	SBR	SBR
Approximation	SBR	SBR	LTI	SBR
Error	LTI	LTI	LTI	LTI
Error Bound	LTI	LTI	LTI	Volitional
Desired Accuracy	Not Answered	Not Answered	Not Answered	Ventriliquation

The remaining four portions of the approximation framework were covered in the prompts in the lab activity. The next one, finding approximations to the unknown value, Leonard was almost able to reason through without the need for further instruction (Table 3). I was called in to Leonard's group to provide additional instruction on what the algebraic representation of an approximation was. After calculating approximations, the lab prompted students to describe the errors of their approximations in the four different representations. This was a task no one in Leonard's group could complete. After asking an undergraduate teaching assistant (UGTA) for help, I was called in to provide further instruction. Leonard did not complete a post-lab that night, but his group members indicated that they had further questions about what errors were and how they were distinct from error bounds. Leonard concurred in his interview: "I didn't really know what to do on this part until Evelyn talked about it the next day in class."

Although Leonard did not complete a post-lab that night, enough students had similar problems on error and error bound, so the intervention the day following the lab activity focused on the definitions of error and error bounds and the difference between these two ideas. All of Leonard's representations on error were coded as learned through instruction.

The concept of the error bound for this lab was only slightly less confusing than error. With additional instruction from an undergraduate teaching assistant, Leonard understood that the error bound was the maximum the error could be, but not how to calculate it. I explained how the error couldn't be any bigger than the distance between the overestimate and the underestimate. When he did, Leonard looked at his graph, and then explained to his group members, that the error bound had to be the vertical distance between the two points they'd already graphed (the overestimate and the underestimate). During the second week of the lab, I helped Leonard write an algebraic representation for the error bound. Hence, all but the graphical context were learned through instruction; since Leonard could not represent an error bound graphically in the next activity, his reasoning here was situation-bound.

The final part of the approximation framework, finding a method to achieve desired accuracy, was the last part of the lab prompt. Leonard's group spent most of the second week on this task. They asked an undergraduate teaching assistant, their instructor, and me for help on how to answer the question. Leonard only answered this question in terms of numerical approximation, but this was the only way any of the students on any lab answered this question. Whether the responses were the result of his spontaneous context-bound reasoning or scientific reasoning taught through instruction, Leonard's final lab write-up was mostly correct with correct reasoning for the context of the lab. Like most students, Leonard had a great deal of difficulty with the next lab.

Lab 4

On Lab 4, the classes struggled so much with the write up that the instructors gave the initial write-ups back ungraded. Leonard was no exception. He did not turn in an initial write-up, and so his first draft would become his final draft. In his interview, he explained that the multiple assignment deadlines that week forced him to prioritize things worth more points.

Leonard's Lab 4 write-up he turned in as final rewrite shows a substantial increase of scientific reasoning. This is an indication that the additional instruction was helpful, but also confirms Leonard's statements that he didn't have a great deal of spontaneous understanding of derivatives, although Leonard required less instruction to complete several tasks successfully on this Lab. This increase in scientific reasoning is evidence students are progressing through their Zone of Proximal Development (Zaretskii, 2009). The brightest spot on Leonard's lab is that he appears to have appropriated one cell of the framework.

Since Leonard never completed the pre-lab, he needed to begin the write-up by completing that assignment. While it was not worth points anymore, he needed to complete these questions in order to be able to do the rest of the parts of the lab. Despite needing to be taught how to complete all of the portions of the pre-lab, the graph Leonard constructed was substantially better than the one in his previous write-up in terms of scale, accuracy, and labeling; it was coded as an increase in quality (Table 4).

Table 4.
Leonard's Lab 4 codes

	Contextual	Graphical	Algebraic	Numerical
Unknown Value	LTI	LTI *Increased Quality*	LTI	LTI
Approximation	LTI	LTI	LTI	LTI
Error	LTI *Less Scaffolding*	LTI *Less Scaffolding*	LTI *Less Scaffolding*	LTI *Less Scaffolding*
Error Bound	LTI	LTI	LTI	APP

Since Leonard stated that he was able to complete a solution for all representations of the approximations after speaking with his instructor and classmate, this portion of his lab was all coded as LTI (Table 4). There was no noticeable increase in the quality of responses than the ones Leonard had in lab 3, but more scaffolding was required for Leonard to complete this portion of the lab. Leonard explained that he received additional, informal instruction from Emily about errors, so this portion of the lab was coded as LTI for each cell (Table 4). However, the amount of scaffolding Leonard needed to complete a solution was substantially less than it was on his Lab 3 write-up; there Leonard needed both scaffolding during the lab and the intervention before he could produce a solution.

Error bound was the part of the lab that Leonard understood the best. After getting help on calculating numerical approximations, Leonard was able to calculate the numerical error bound on his own, so this cell of the approximation framework is appropriated. “After I had the error bound in numbers, I wasn’t sure what to do for the other parts, so I asked [my instructor”, Leonard said. Although he was very comfortable with calculating numerical error bounds, Leonard had not internalized what error bounds were in non-numerical contexts. Generally, numerical error bound was the first cell any student appropriated (Table 4)

For the final question about how to approximate within any error bound, Leonard neither received additional instruction nor completed the task volitionally. “I took the answer in Lab 3 and I copied it down here. I lost some points last time, so I tried to say something about Iodine, like Emily did,” Leonard explained. Both the contextual and the numerical answers he gave were coded as ventriliquation.

Lab 7

Leonard did not complete a Lab 7 write-up. Given the lack of documents to code, I asked Leonard to complete Lab 7 as part of a task based interview, so that I could gather more information about the depth of his understanding about approximation in the context of definite integrals. For this portion of the interview, I asked Leonard to answer the questions in the lab orally. He missed both days following the lab this unit where his instructor provided additional scaffolding, so it was unlikely that he received additional instruction in a formal setting. After each response, I asked Leonard why he gave the answer he did, and if anyone had helped him understand that part of the lab. In terms of the approximation framework, this was the lab Leonard felt he completed with the least help and understood the most.

The first part of the lab Leonard and I talked about was the unknown value portion of the approximation framework. As usual, this piece had been to the students to compete as a pre-lab. Leonard explained that his group member Emily explained to him after my demonstration with a rubber band why calculus was needed, and she helped him with the algebra. Based on Leonard’s response, I concluded that creating his graphical and numerical representation of the unknown value were volitional acts, but that he learned what the unknown value was through my scaffolding, and how to represent it algebraically through Emily’s help. I coded both of those cells of the approximation framework as LTI (Table 5).

Table 5
Leonard’s Lab 7 codes

	Contextual	Graphical	Algebraic	Numerical
Unknown Value	LTI	VOL	LTI	VOL
Approximation	VOL	VOL	LTI	VOL; LTI
Error	VOL	VOL	LTI	VOL
Error Bound	VOL	VOL	LTI	APP
Desired Accuracy	VOL	VOL	Not Answered	VOL

I coded the contextual and graphical representations as volitional. The numerical representation was both volitional when Leonard could do the calculations by hand and learned through instruction when the calculator handout was needed. The algebraic representation of a Riemann sum was also coded as learned through instruction, although it is clear that Leonard does not have complete mastery of that concept at this time (Table 5).

Error was one part of the approximation framework that Leonard struggled with throughout the semester, but he was able, within this context to reason and explain errors to me in the course of the interview. Since Leonard was able to easily explain three of the representations, I coded the contextual, graphical, and numerical representations as volitional, and the algebraic representation as learned through instruction (Table 5). This was typical of Leonard; he had some of the weakest algebra skills in the class and often needed help with the symbolic manipulation.

Error bound has been the area of the approximation framework that Leonard was most comfortable with and this lab was no exception The contextual and graphical representations were coded as volitional (Table 5), while the numerical representation was coded as appropriated based on his past performance. I coded the algebraic context as learned through instruction, because Leonard needed help to assemble the components of the error bound; however, he was able to write down the algebraic representation volitionally once he had that assistance.

For the final portion of the lab, Leonard freely admitted that on the first two labs, he just transcribed what someone else told him to say, but this time he had more ownership this concept:

> If we want to make sure we approximate the work to within some number of sigfigs, we have to first calculate how many rectangles to use. Then we can't really draw of a graph of that because there are usually too many to draw. Then you find the left and right Riemann sums. As long as you calculated right, everything should work out.

After taking a detailed look at Leonard's progression throughout the semester, I wondered how he compared to the other students who earned similar final grades. Sandra, who seemed so similar at first glance, had a different learning trajectory due to her difference in participation and lower starting point.

Sandra

Sandra was conditionally admitted to a MA program in bio-chemistry, where the successful completion of introductory calculus was one of the requirements she needed to meet. Although she did not have a pre-calculus course and was several years removed from her last formal mathematics course, she was motivated to succeed and willing to ask for help on behalf of her table when she was confused. Sandra explained in her first interview that she understood the basic idea, but had trouble articulating what she knew into a coherent solution:

> I had the general idea. We couldn't factor to find where the hole in the graph was, so we had to plug in points on either side to get an idea where it was. I got that. Where I had a lot of problems on this lab was understanding how to write things down in a way that made sense to [my instructor].

Lab 3

On the first approximation lab, Sandra spent extensive time receiving help from the UGTA's and her instructor, both in class and in office hours. Her write-up, unlike the other students, shows no situation-bound reasoning codes, and only one spontaneous reasoning code, on this lab. The rest was LTI, ventriliquation, and volitional on the numerical parts of the lab (Table 6).

Table 6
Sandra's Lab 3 codes

	Contextual	Graphical	Algebraic	Numerical
Unknown	Not Answered	UH	VOL	SBR

Value				
Approximation	LTI	LTI	LTI	VOL
Error	LTI	LTI	LTI	LTI
Error Bound	LTI	LTI	LTI	VOL
Desired Accuracy	LTI	Not Answered	Not Answered	Not Answered

Even on this first lab, Sandra and Leonard differ on when and how they seek help. If Leonard needed help, he would ask his group member Emily; Leonard did not consider the post-lab as a possible source of assistance. Sandra, who was several years older than her classmates, did not consider her group members as a viable source of help. Sandra was willing to ask the UGTA's questions, but preferred to ask her instructor if she could. Where Leonard was able to use intuition and spontaneous reasoning to construct his solution to this lab; Sandra's algebra was rusty and she needed review on the background information before she could begin the lab. Since Sandra also received extensive help on how to write the solution, Sandra's reasoning on Lab 3 was scientific rather than spontaneous.

Sandra completed both of the post-labs associated with this first approximation framework lab. On the first, she was able to accurately how far her group got, and where she was stuck – on the difference between errors and error bounds. During the next, class, this was the topic her instructor discussed. On the final post-lab, Sandra observed that the method to solving all of the contexts was identical to the one she used, which was the connection we wanted students to make after the Jigsaw. At the end of the post-lab, Sandra wrote what became her default response at the end of every post-lab: I don't have any questions on the lab at the moment, but I will come and find you if I have any. In her interview, Sandra explained why she stopped asking questions on the post-labs:

> The one time I asked a question, [my instructor] covered it in class. Since I'm towards the bottom of the class, there couldn't have been too many others with the same question. I don't want to slow the class down, but this way I can let [my instructor] know that I need help without having to say anything in front of the kids. I know [my instructor] cares about my success because he talks about what I say on the pre-labs and post-labs, so I knew I can go to office hours and not be judged there.

Sandra greatly overestimated the abilities of her classmates; the difference between error and error bound was the question that every student that had a question asked. Both her and in her final interview, Sandra mentioned the importance of being able to let her instructor know that she needed and would seek help without having to ask questions in front of her younger classmates.

Lab 4

Like Leonard, Sandra completed very little of the first Lab 4 write-up. Sandra appropriated the numerical unknown value, volitionally identified the unknown value's context, and had a partially correct graph coded as PRW. Like Leonard, she didn't turn in a write-up and the multiple deadlines that week in calculus as the reason she didn't have time to do more on the lab. Sandra's second attempt at this lab, where she got extra help in and outside of class from her instructor, was able to complete far more of the assignment. Like Leonard, most of this lab was coded as learned through instruction. However, some of the graphical parts were situation-bound reasoning, so there is some evidence of spontaneous understanding at the end of this activity (Table 7).

Table 7

Sandra's Lab 4 codes

	Contextual	Graphical	Algebraic	Numerical
Unknown Value	SBR	INQ	APP	PRW
Approximation	LTI	SBR	LTI	SBR
Error	LS	VOL	LTI	LS
Error Bound	LTI	LTI	LTI	APP
Desired Accuracy	INQ	Not Answered	Not Answered	LS

On this lab, Sandra sought less help than on the first lab. Unlike Leonard, Sandra used spontaneous reasoning in the new context of derivatives, and only asked for assistance once she was stuck. Although Sandra correctly identified the algebraic representation of the unknown value ($f'(2)$) in her pre-lab, her initial approximations were y-values rather than slopes. Sandra got help from an UGTA on what the approximations and algebraic representations were for a derivative. For error bounds, Sandra visited her instructor outside of class for all but the numerical representation that she appropriated. For all students in the study, numerical error bound was the first cell appropriated. The general progression on this lab from lab 3 is that while Sandra still sought help for almost every part of the lab, she attempted more cells on her own and required less help to complete the solution.

The scaffolding Sandra received was more formal than the scaffolding Leonard sought. Where Leonard would ask his group members for help if he had trouble, Sandra went to her instructor. Sandra also asked much more targeted questions; where Leonard would ask for help on a row of the approximation framework, Sandra asked specific questions for each cell.

Sandra did not complete any of the post-labs for this lab. She cited two reasons for this:

> It wasn't that I didn't want to, but this past three weeks has been awful. In [Organic Chemistry], we had our midterm. In here we had a lab, the gateway, and a chapter test. Something had to give, and this was the thing that would hurt me the least. Besides, I know I can go to [my instructor] if I really needed to, and everyone else must have the same questions I do, because what I need help on the most is usually what got covered in class.

Lab 7

Sandra's final approximation lab of the semester was her best lab. She needed the least amount of help to complete her write-up, and the only area that was still learned through instruction were all of the representations of error and help with summation notation. Most of Sandra's lab was completed volitionally, and she appropriated an additional portion of the approximation framework (Table 8).

Table 8
Sandra's Lab 7 codes

	Contextual	Graphical	Algebraic	Numerical
Unknown Value	INQ	APP	APP	APP
Approximation	VOL	Not Answered [VOL]	LS	VOL
Error	APP	APP	LS	APP
Error Bound	LS	LS	LTI	APP
Desired	SBR	SBR	SBR	SBR

In Lab 4, Sandra's solution showed discrete instances of scientific or spontaneous reasoning. In the final approximation lab, most cells show her integrating her intuition with the instruction given in class. The scaffolding Sandra received on this lab was on the level of a reminder of something done in a previous class rather than explicit instruction on how to solve parts of the lab; with the exception of the notation needed for the error bound which was covered in class. Although, Sandra did not answer the graphical approximation question on her lab, when I asked her about it in her interview she admitted that she didn't see that part of the question and immediately labeled the approximation correctly on a copy of her lab.

Sandra shows a slightly atypical pattern of appropriation; while the numerical representation and unknown value were the two areas of the approximation framework appropriated first and second, every other student appropriated approximations and error bound before errors. Sandra explained why in her second interview: "Errors were the hardest part of these labs for me to understand. I spent a lot of time with [my instructor] on them all semester. When we got to this lab, everything just made sense."

Sandra completed both post-labs for Lab 7. On the first post-lab, she answered the calculation question without any errors and explained that her group was about half done with their lab write-up. Although they were having a little bit of trouble on the calculator, everything else was going smoothly. The final post-lab, completed after Lab 7 was turned in, asked students to summarize what they learned in the semester during the labs, and asked them to define what a limit was in their own words. Sandra's response is given below. She shows an understanding of most parts of the framework, and defined a limit, albeit in an odd way, in terms of approximation:

> Calculus is what you do when you can't solve a problem with algebra. First, you have to figure out something you can calculate with algebra that will approximate what you want to know. Since the algebra isn't quite right, you get errors, but you can't calculate them, because you don't know the real value. So the error bound, overestimate – underestimate, is how off you can be. If we have to, we get within any error, but it will probably be a real pain to calculate. We haven't talked about what a limit is in a long time, but the limit is the real answer, the perfect approximation.

In a separate educational ethnography following up this dissertation project (Goss & Dibbs, 2013), Sandra, the only interview participant to earn a C and continue on to the second semester of calculus, continued this pattern of steady improvement, and appropriated the entire approximation framework. Sandra improved her grade steadily throughout the semester, from a mid-D to a mid-C, and continued that trajectory after this semester. She earned an A- in Calculus II. For much of her work, it appears that being rusty of the prerequisite knowledge was the primary obstacle to her success during the first semester.

Discussion

Although Sandra passed the class and Leonard did not, most the difference in their grades can be accounted for by the fact that Leonard did two post-labs and failing to turn in Lab 7 and Sandra did all but two. While one could argue that Sandra was a better student, Sandra actually had a lower grade than Leonard did until the beginning of the fifth lab and did not find her group members a source of support. Given that, which functions of formative assessment were most prominent in Sandra's conceptual development of the approximation framework?

None of the nine students interviewed in this study ever mentioned formative assessment in terms of the first purpose, clarifying learning intentions. They saw the labs as extra practice with the big ideas of calculus covered that week. The one thing that was mentioned as an important tool for clarifying learning intentions of the instructor was the set of problems posted to the course webpage about a week before each test. There is no data to indicate that this purpose of formative assessment was relevant to students' conceptual development of the approximation framework.

Both Leonard and Sandra mentioned that the intervention instruction the day after the labs as a crucial part of the lab; both said that they would have had a much harder time completing the lab write-up without this additional instruction. Leonard went so far as to cite the effectiveness of the next day intervention as the main reason he never completed post labs; he claimed all of his questions got answered the next day in class without the effort of completing the post-lab. The instructors designed the intervention around a summary report of their class' formative assessments; I gave each instructor a list of the three most commonly asked questions and one question that would be nice to answer if they had time. The formative assessment in the form of the post-lab certainly helped instructors design instructional activities that students felt was important to moving them forward on their understanding of the lab.

Neither Leonard nor Sandra ever asked a question on the post-lab that was not covered in class during an intervention, so providing individual feedback that moved learners forward was not relevant to their conceptual development. However, this was not the case for all participants in the larger study. The data collection precluded the pre-lab for being a means for students to become learning resources for each other, except in the sense that all of the groups started working on the lab quickly, because some student in the group had a plan on how to create approximations.

None of the participants mentioned affective variables or metacognition skills improving, even with probing, as part of the benefit of the formative pre-labs and post-labs that students completed. Sandra's quote that mentioned that formative assessment was evidence that the instructor cared about their learning was a prevalent theme, and she said that she tried harder as a result of this caring. She also was able to ask much more specific questions about her difficulties in the lab, which does suggest some improvement in metacognitive skills. The most important benefit for the pre-labs and post-labs for Sandra was the chance to peripherally participate in the class by asking questions directly to her instructor on the post-labs or signaling that she needed more help on the lab without having to admit her need in front of her classmates. This opportunity for peripheral participation and the targeted instruction were the two most important purposes of formative assessment for Sandra.

Overall, the peripheral participation and targeted instruction allowed students develop a deeper conceptual understanding of the approximation framework with an efficient use of class time. The interventions for all seven labs accounted for just over 1.5 class periods, spread over the fifteen week course. One potential pitfall of formative assessment in general was echoed by Sandra: peripheral participation and formative assessment is only perceived as effective to students when it is clearly signaled to students that the instructor read and used the formative assessments in some meaningful way. The act of giving students a formative assessment will not have measurable benefits without visible use to the students (Black & Wiliam, 2009).

References

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

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

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

Goss, M., & Dibbs, R. A. (January, 2013). *Group norms establishment and enculturation into a calculus II classroom using the approximation framework*. Unpublished paper presented at the Joint Mathematics Meetings, San Diego, CA.

Lave, J., E. Wenger. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge England: New York.

Merriam, S.B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jonsey-Bass.

Oehrtman, M. (2008a). "Layers of abstraction: Theory and design for the instruction of limit concepts". In *Making the connection: Research and practice in undergraduate mathematics, MAA Notes Volume*, Edited by: Carlson, M. and Rasmussen, C. 65–80. Washington, DC: Mathematical Association of America. 73.

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

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, California: SAGE Publications, Inc.

Smagorinsky, P. (1995). The social construction of data: Methodological problems of investigating learning in the Zone of Proximal Development. *Review of Educational Research, 65*(3), 191-212.

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

CALCULUS STUDENTS' UNDERSTANDING OF VOLUME

Allison J. Dorko Natasha M. Speer
Oregon State University University of Maine

Researchers have documented difficulties that elementary school students have in understanding volume. Despite its importance in higher mathematics, we know little about college students' understanding of volume. This study investigated calculus students' understanding of volume. Clinical interview transcripts and written responses to volume problems were analyzed. One finding is that some calculus students, when asked to find volume, find surface area instead and others blend volume and surface area ideas. We categorize students' formulae according to their volume and surface area elements. We found that some of these students believe adding the areas of an object's faces measures three-dimensional space. Findings from interviews also revealed that understanding volume as an array of cubes is connected to successfully solving volume problems. This finding and others are compared to what has been documented for elementary school studens. Implications for calculus teaching and learning are discussed.

Key words: student thinking, calculus, volume, surface area

Introduction

This paper details calculus students' understanding of volume in non-calculus contexts. Many calculus topics involve volume, including optimization and related rates in differential calculus, volumes of solids of revolution and work problems in integral calculus, and multiple integration in vector calculus. Although volume shows up in these places and others throughout mathematics curricula (e.g., geometry, word problems), researchers have focused on elementary school students' difficulties. Less is known about how calculus students understand volume.

The purpose of this study was to investigate what calculus students understand about volume in *non-calculus* contexts. A non-calculus focus was intentional: in other areas of mathematics, researchers have found that students' prior knowledge interacts with their learning of calculus. Two notable examples of this are student understanding of function (Carlsen, 1998) and variable (Trigueros & Ursini, 2003). Findings indicate that what appear to be difficulties in learning and understanding calculus may in large part be derived from difficulties with these underlying, non-calculus concepts. Therefore, investigating calculus students' understanding of volume in non-calculus contexts sheds light on student difficulties with volume-related calculus ideas and builds a foundation for studying student difficulties with calculus topics that use volume, such as those noted above.

We know from research that volume presents challenges to elementary school students (Battista & Clements, 1998; De Corte, Verschaffel, & Van Collie, 1998; Fuys, Geddes, &Tischler, 1998; Hirstein, Lamb, & Osborne, 1978; Iszák, 2005; Lehrer, 2003; Lehrer, Jenkins, & Osana, 1998; Mack, 2011; Nesher, 1992; Peled & Nesher, 1998; Simon & Blume, 1994). These difficulties are also reflected in student performance on standardized test items. For example, on an eighth grade NEAP multiple-choice question (U.S. Department of Education, 2007), students were given the dimensions of five rectangular prisms and asked which had the greatest volume. Only 75% of students answered correctly, which indicates that students may enter high school (where instruction builds on presumed competence with volume concepts) without proficiency in volume calculations. It would be useful to know if the difficulties elementary school students face persist through high school and into their study of college-level mathematics.

We conducted this study within a cognitivist framework (Byrnes, 2001), giving students mathematical tasks and analyzing the reasoning underlying their answers. This is consistent with the cognitivist orientation toward focusing on "the cognitive events that subtend or cause behaviors (e.g., [a student's] conceptual understanding of the question)" (Byrnes, 2001, p.3). We collected written survey data and conducted clinical interviews to investigate the following research questions:

1. How successful are calculus students at volume computational problems?
2. Do calculus students find surface area when directed to find volume?

Our major finding is that nearly all students correctly calculate the volume of a rectangular prism, but many students perform surface area calculations or calculations that combine volume and surface area elements when asked to find the volume of other shapes.

Student Thinking about Volume

Volume is first learned in elementary school (NCTM, 2000; NGA & CCSSO, 2010) and, as noted above, little literature exists about calculus students' understanding of volume. Key issues that have been the focus of research include elementary school students' understanding of arrays and area and volume formulae as well as secondary school students' understanding of cross-sections.

Elementary School Students' Volume Understanding

Volume computations rely on the idea of array of cubes. A three-dimensional array is formed by the iteration of a cube into rows, columns, and layers such that there are no gaps or overlaps. Two difficulties students have are understanding an array's unit structure (Battista & Clements, 1996) and using an array to compute volume (Curry & Outhred, 2005).

These are related difficulties. One source of difficulty with using an array for computation is not seeing the relationships between rows, columns, and layers. Some students, given an array of cubes and asked to find volume, counted individualized cubes with "no global organizational schema" and seemed to view the answer as representing "a large number of randomly arranged objects" rather than a count that represented the array's volume (Battista & Clements, 1998, p. 228-229). Other researchers have concluded that elementary school students seem to see units as individual pieces to count rather than fractional parts of an initial whole (Hirstein, Lamb, & Osborne, 1978; Mack, 2011). Students who counted individual cubes neglected the innermost cubes and sometimes double-count edge and corner cubes (Battista & Clements, 1996).

Battista and Clements (1996) studied students' enumeration of three-dimensional cube arrays using written and manipulative tasks and found that only 23% of third graders and 63% of fifth graders could determine the number of cubes in a 3x4x5 cube building made from interlocking centimeter cubes. The researchers concluded that "students might see the three-dimensional array strictly in terms of its faces" (Battista & Clements, 1998, p. 229). In other words, these students may have been thinking about surface area when asked about volume.

Curry and Outhred (2005) found that students who are successful at enumerating arrays of cubes seem to have a mental picture of arrays and use a computational strategy of counting the units in the base layer and multiplying by the number of layers (Curry & Outhred, 2006) while the unsuccessful students typically covered only the base of the box. The researchers concluded that although "most students seem to have achieved a sound understanding of length and area measurement by Grade 4, the same cannot be said for volume [arrays]" (p. 272).

Some elementary school students use area and volume formulae without understanding them (De Corte, Verschaffel, & van Collie, 1998; Fuys, Geddes, & Tischler, 1988; Nesher, 1992; Peled & Nesher, 1988). In a study about students' multiplication strategies, De Corte et al. (1998) included area computation problems and found that students may multiply length times width to find area "not [as] a result of a 'deep' understanding of the problem structure and a mindful matching of that understanding with a formal arithmetical operation, but… based on the direct and rather mindless application of a well-known formula" (p. 19). Echoing this, Battista and Clements (1998) found some students use V=LWH "with no indication that they understand it in terms of layers" (Battista & Clements, 1998, p. 222). Even some prospective elementary school teachers use the A=lw formula without being able to explain why it finds area (Simon & Blume, 1994).

Secondary school students' understanding of cross-sections

Identifying the shape of a solid's cross-section is difficult for middle school and high school students (Davis, 1973). This finding is important because some volumes can be thought of as V=Bh where B is the area of the base of the solid and the base is, in fact, a cross-section. This finding carries particular importance if it is also true for calculus students, as volumes of solids of revolution problems require identifying the shape of a cross-section.

The present study was designed to investigate calculus students' computations of volume, their understanding of volume formulae, as well as the issues noted above that other researchers have documented in younger students.

Research Design

Data Sources and Instrument

The data analyzed here are from written surveys completed by 198 differential calculus students and 20 clinical interviews with a subset of those students. Subjects were enrolled in differential calculus at a large public northeastern university and the researchers recruited volunteers to complete written surveys and clinical interviews. Data were collected for three semesters: spring 2011, summer 2011, and fall 2012. The university offers a single track of calculus for all majors in the physical sciences, engineering, biological sciences and education, as well as other disciplines.

Data collection and analysis had two phases. First, students completed written tasks. Since our focus was on the reasoning behind students' answers, we interviewed a subset of the students so that we could hear how students reasoned through the problems and ask questions about their reasoning. This methodology allowed for a quantitative analysis of a large number of written responses and a qualitative analysis of student thinking about those responses.

The written survey tasks consisted of diagrams of solids with dimensions labeled. Students were directed to compute the volume of the solid and explain their work. The rectangular prism task is shown in Figure 1. The other tasks were:

- A right triangular prism; triangular base l= 3 ft, h = 4 ft; h_{prism}= 8 ft
- A cylinder, r = 3 in., h = 8 in.

The complete statements of these tasks can be found in Appendix A.

Interviewees completed the written instrument but were asked to "think aloud" as they worked on the tasks. Clarifying questions were asked to probe understanding. Commonly-asked questions of this sort were "Can you tell me about that formula? Why is the 2 there?" and "Can you tell me why that formula finds volume?" Interviews were audiorecorded and transcribed.

What is the volume of the box? Explain how you found it.

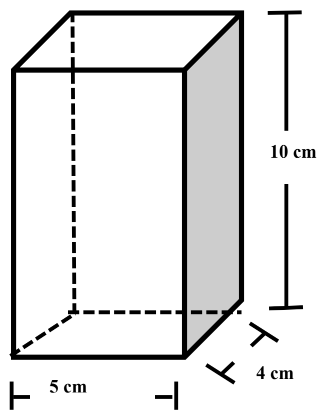


Figure 1. Volume of a Rectangular Prism

Method of Data Analysis

Data were analyzed using a Grounded Theory-inspired approach (Glaser & Strauss, 1967). This entailed looking for patterns in a portion of the data and forming categories, then creating category descriptions and criteria. Those criteria were then used to code all the data, refining categories until new categories ceased to emerge. One departure from classic Grounded Theory was accessing literature prior to coding. A second was the use of anecdotal evidence that calculus students sometimes find surface area when asked to find volume. These departures informed coding in that prior to looking at data, we had ideas about what categories might emerge.

Analyzing written responses required deciding which parts of a response were relevant to the research question. We used the magnitude of the answer to judge correctness and we looked at written work (arithmetic) as it gave clues to student thinking. Units were not taken into account here, though the units students use for spatial computations is an additional issue that was part of a larger study (Dorko, 2012).

Analysis was done by shape, not by student. That is, the data presented are the percent of students whose responses fell into each category for that task. The initial analysis resulted in three categories for students' work: volume, surface area [instead of volume], and other. The categories and their criteria are presented in Table 1. We used these criteria to develop coding algorithms for the three volume problems. An example of a task and its algorithm are shown in Figure 2. All algorithms were created in a way that sorted responses based on identifying parts of a student's work that might represent finding surface area, parts of a student's work that might represent finding volume, and an "other" cateogry for responses that had neither of the aforementioned ideas. That is, algorithms for the other shapes are similar to the algorithm presented below.

Category	Found Volume	Found Surface Area Instead of Volume	Other
Criteria	Magnitude is the correct magnitude of the object's volume, or magnitude is incorrect for the object's volume but the work/explanation is consistent with volume-finding (i.e., multiplication or appropriate addition)	Magnitude is the magnitude of the object's surface area or the student work/explanation contains evidence of surface-area like computations, such as addition. To allow for computational errors, magnitude may or may not be the actual magnitude of surface area.	Student found neither volume nor surface area

Table 1. Categories for written responses.

What is the volume of the cylinder? Explain how you found it.

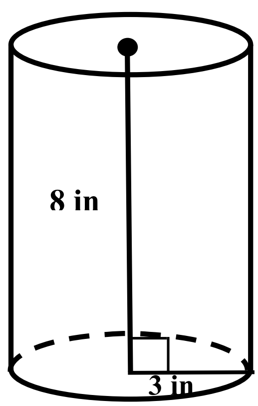


Figure 2. Volume of Cylinder task.

Coding algorithm for the cylinder (correct volume: 72π units^3; correct surface area: 66π units^2)

1. If the work says $\pi r^2 h$, $2\pi r^2 h$, 72π, or 144π, categorize as "found volume." If not, go to step 2.
2. Did the student write πr^2 + ________ or $2\pi r^2$ + ____________where ________ is something that looks like it might be πdh or some other computation that looks like an area of a lateral face? Did the student write 66π? In either case, categorize as "found surface area instead of volume." If not, proceed to step 3.
3. Categorize as "other."

The method of analysis for interview data mirrored the method of analysis for survey data. As interview data included both transcripts and students' written work, there were two parts to the analysis. First, written data were categorized according to the aforementioned algorithms. Then, transcripts were used to investigate the thinking that led to answers for each category. For instance, we looked for students who used the formula $2\pi r^2 h$ and asked the student to "unpack" the formula. Specifically, we looked for an explanation of why the 2 was there. Did the student think the area of a circle was $2\pi r^2$? Did the two refer to the two bases of a cylinder? A yes to the first would be consistent with thinking of volume as *area of base times height*, albeit with an incorrect formula for the area of the base. A yes to the second would be consistent with the surface area idea of including the areas of all faces. This analysis was used in two ways: to sort student formulae and to investigate why students find surface area instead of volume.

In the next section, we present the findings from our analyses. We use interview to further explain some of these results. In particular, we use interview quotes to clarify how we classified students' volume formulae. The second section of the results, students' success rates on the tasks, relying on coding from the written data. We return to the interview data to discuss why some students find surface area and the relationship between students' understanding of arrays and their computational success.

Results and Discussion

We present our findings in four parts. One finding is about students' reasoning and computational formulae, and since this is a good overview of the issues students have with volume, we begin with that and follow it with success rates on the problems. We then discuss why some students find surface area instead of volume. The final section is about the relationship between success on problems and understanding arrays.

Students' Volume Formulae

We believe there is an important link between students' formulae and their reasoning: that is, our data leads us to believe that students' formula are not (as is commonly assumed) remembered or misremembered, but are instead representative of ideas students have about volume. This finding, based on the synthesis of interview data with written work, led us to categorize students' formulae according to their surface area and volume elements. What we mean by "surface area and volume elements" is what we alluded to in discussing how the 2 in $2\pi r^2h$ might be from an ill-remembered area formula and might be from accounting for two bases. Categorizing students' formula in this way gave us the categories and component formulae shown in Table 3. The example given is for the cylinder; similar tables exist for each shape and are included in Appendix C. (Note the appearance of $2\pi r^2h$ in both the "incorrect volume, no surface area element" and "surface area and volume elements" categories, per the reasoning stated above).

Correct volume	Incorrect volume, no surface area element	Surface area and volume elements	Surface area	Perimeter
πr^2h	$2\pi r^2h$ $(1/3)\pi r^2h$ $(1/2)\pi r^2h$ $(4/3)\pi r^2h$ πrh $(1/2)\pi rh$ $h*d*r$	$2\pi r^2h$ $2\pi rh$ $2\pi r + \pi rh$ $\pi r^2 + 2\pi d$ $2\pi r^2 + 2rh$	$2\pi r^2h + 2\pi rh$ $2\pi r^2h + \pi dh$	$d+h$

Table 2. Categories for student responses to the cylinder task

This table includes all formulae that appeared in students' written work and interviews. Interview data provided help in placing the formula, and interview data are the basis of our claim that students' formulae are a reflection of their reasoning. For instance, consider Nell's reasoning about the volume of the cylinder:

> **Nell:** I don't know the formula for this one. Two pi r squared… times the height. Sure. We'll go with that one. So you have two circles at the ends, which is two pi r squared… you have two pi r squared because that's the area on the top and the bottom so you can just double it, then you have to times it by the height.
> **Interviewer:** Why do I have two areas?
> **Nell:** You have two circles.
> **Interviewer:** What about this multiplying by the height? Why do we do that?
> **Nell:** It gives you the space between the two areas. Volume is all about the space something takes up so you need to know how tall it is.

Nell's reference to the space between two areas is indicative that she was thinking about volume. However, her formula ($2\pi r^2h$) included a surface area idea: she explained it as "the area on the top and the area on the bottom, so you can just double [πr^2]". We thus put the formula $2\pi r^2h$ in the "surface area and volume elements" category (see Table 3). It is also included in the "incorrect volume, no surface area elements" because other students talked about this formula as *area of base times height* where the area of the base was $2\pi r^2$. In this case, the two is not a nod to two bases, it is an incorrect formula for area but correct reasoning for volume.

Nell was not the only student who thought about including both circles when finding volume: Jo went back and forth about whether she should use the formula $2\pi r^2h$ or πr^2h. The interviewer asked her to make the case for both one and two circles as a way to investigate her reasoning:

> **Jo:** The area of the circle is pi r squared times the height, but I can't decide if I need one or two circles.
> **Interviewer:** Convince me that you need two circles.
> **Jo:** You need two because you have the top and the bottom of the cylinder. But you don't actually need two… you just need the one. Because you get the area of the circle and you multiply it by the height… the circle is the same throughout the whole layer so you just multiply it by the height.

Jo's final reasoning was correct, but it's noteworthy that her initial response to the problem involved a surface area idea. Thus, despite her correct final response, we believe this is evidence that some students have mixed and combined surface area and volume ideas.

Most of the elements of the categories shown in Table 3 come from students' written work. No interviewee used a formula like $(4/3)\pi r^2h$ (or any of the others with a fractional coefficient for an otherwise correct volume formula), but we suspect students mixed and combined the formula for a cylinder with that of a sphere, pyramid, or cone – all shapes whose volume formulae have fractional coefficients. Further, we suspect that students who use these formulae do not have an understanding of volume as *area of base times height*. Our evidence for this claim is that for a student who understands volume as area of base times height, a formula like $(1/3)\pi r^2h$ makes little sense.

The other formulae in the table provide additional evidence that some calculus students have difficulties with volume and surface area. We think answers like h*d*r and πrh (both from "incorrect volume, no surface area elements"), in which it seems the student has multiplied whatever dimensions were given (and in the latter, probably remembered that circle calculations often involve π), may result from translating a V=lwh form to a different shape. That is, we speculate that the students have a schema for volume to the effect of "volume is the product of measured attributes" and roughly equated V=lwh to V=hdr. Other instances of this included multiplying all of the dimensions given in the triangular prism; for instance, students' volume calculations included formulae like and 3*4*5*8, (1/2)*3*4*5*8. These lend further evidence that some students may hold "multiply whatever numbers you're given" as an accurate way to find volume, and moreover, don't understand volume as *area of base times height.*

In conclusion, we found that calculus students who are unsuccessful at finding volume often find surface area or a number that represents a combination of surface area and volume ideas. We speculate that many students do not understand volume as area of base times height, and construct formula based on ideas about area and volume. Some of those ideas include volume formula often having fractional coefficients (as in the (1/2)*3*4*5*8 and $(4/3)\pi r^2h$ cases), or the more troublesome cases in which students have combined surface area and volume ideas. In the next section, we discuss the prevalence of these sorts of difficulties.

Students' performance on volume tasks

The counts and percentages of students who found volume, surface area, or other for the four solids are shown in Table 4.

	Rectangular prism (n=198)	Cylinder (n=198)	Triangular Prism (n=122)
Found Vol.	194 (98.0%)	172 (86.9%)	95 (77.9%)
Found SA	3 (1.52 %)	10 (5.1%)	17 (13.9%)
Other	1 (0.5%)	16 (8.0%)	10 (8.2%)

Table 4. Counts and percentages of students who did and did not find volume

This table shows that 98% of students found the volume of the rectangular prism; 87% of students found the volume of the cylinder, and 78% of students found the volume of the triangular prism. We speculate that a rectangular prism is a more common shape and thus students have its volume formula memorized, but struggle when they encounter other shapes. Certainly, Nell and Jo struggled with how they might find the volume of the cylinder. We also note that students were more likely to find surface area for the cylinder and the triangular prism (5.1% and 13.9%, respectively). Further research is needed to explain exactly why different shapes are harder in terms of finding volume, but we suspect the reason may be students not understanding volume as *area of base times height* or the combination some students hold of surface area and volume ideas.

Why some students find surface area

We used interview data to examine why some students find surface area when directed to find volume. One reason is that some students seem to have combined elements from formulae, as Nell and Jo did. Despite the fact that both of these students understood volume as three-dimensional space, as evidenced by Nell's comment about volume being the 'space between the two areas' and Jo's comment that the circle was "the same all the way through so we can just multiply by the height." A different reason that some students find surface area is the belief that *adding the areas of the faces measures three-dimensional space.* For instance, Geddy's description of filling an object with sugar cubes is indicative of understanding the concept of volume:

Geddy: Volume is the amount of units it takes to occupy a space, like a three dimensional space. If you think of a box of sugar cubes, like a Domino box, I think when they come packaged they are usually just full of the little sugar cubes and there's no space between those cubes. So that's what volume is. It's when you have a bunch of little smaller pieces combining to fill a space without any gaps.

Geddy's "volume" computation, however, was actually a surface area computation (see Figure 5).

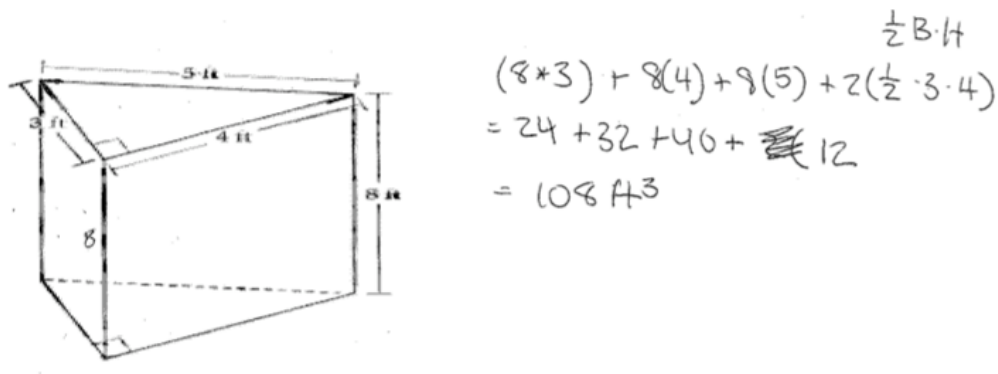


Figure 5. Surface area of the triangular prism

Alex also understood volume but found surface area. She described volume as "when I think of volume I think of, like, this water bottle – what's the volume of water it can hold." Talking about holding water is evidence of understanding volume as three-dimensional space. However, Alex too found the surface area of the triangular prism. Her work is shown in Figure 6.

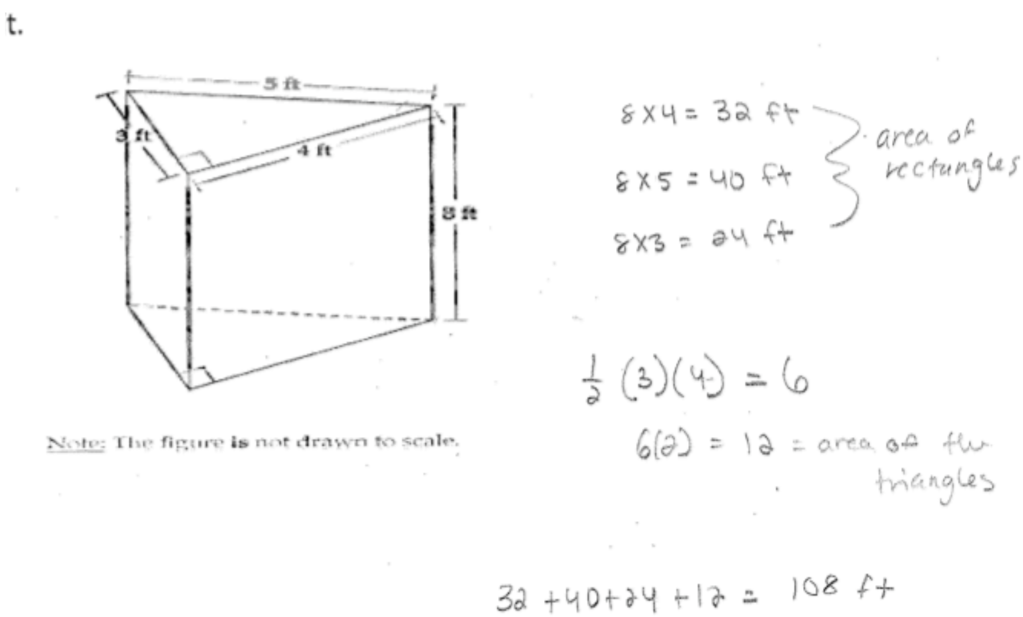


Figure 6. Alex's triangular prism

We asked Alex to explain her work.

Alex: I took the area of each rectangle and added that up, then I took the area of the triangles and added that to the rectangles to get the overall area. And I couldn't remember the area for the triangle. I thought it was ½ times base times height, which is 6. And there are 2, so 6 times 2 equals 12, so 12 is the area of the triangles… and then the area of the rectangles… and I just added them all together.

There is a discrepancy in Alex, Geddy, and other students' understanding of volume as a concept and their calculations such that these students understand volume, but think adding the areas of the faces accounts for the measure of a three-dimensional space. This reasoning, and the combination of surface-area-and-volume formula discussed above, are the two reasons that students in this study found surface area when directed to find volume.

An understanding that seems to be connected to students' success on area and volume tasks is their understanding of arrays. We discuss this finding in the next section.

Students' Understanding of Arrays

A number of students used the formula V=lwh to find the volume of the rectangular prism. We asked interviewees to "unpack" this formula to see if they were simply reciting a formula or if they understood why it finds volume. Students' responses led us to several findings about their understanding of arrays. In this section, we discuss these findings and compare them to findings about elementary school students' understanding of arrays.

One finding about calculus students is that some students can describe the formula V=lwh in terms of relationships between rows, columns, and layers for a rectangular prism but not for other shapes. For instance, Amelia used the LWH formula for the volume of the rectangular prism and talked about an array when the researcher asked her to unpack that formula:

Amelia: If we think about this in terms of area – you still have like this box [points to the 5 cm x 10 cm face], as long as you can figure out that there's like [draws a 5 x 10 array of swuares on the face] so this represents 50 boxes, then you know that you have four of these… so you can think of it as having four sheets of 50 squares.

Amelia's drawing is shown in Figure 7. She has drawn 'boxes' (unit squares) on the front face and indicated the four 'sheets' (layers) along the 4 cm orthogonal face.

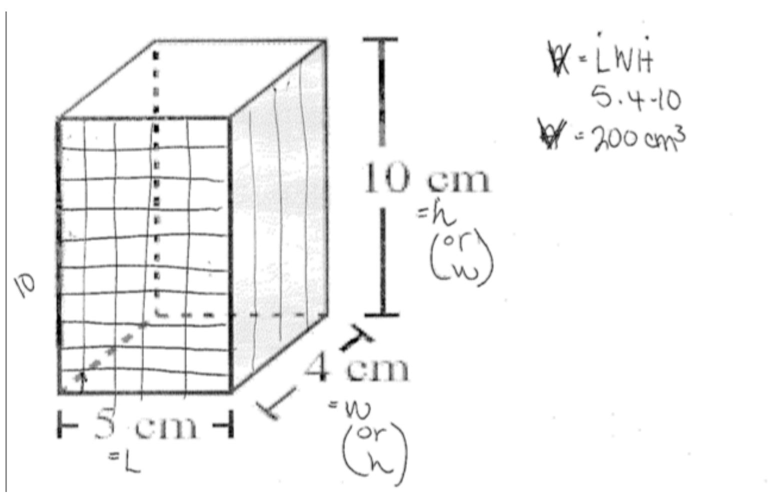


Figure 7. Amelia's rectangular prism

While this work is indicative that Amelia understands arrays, her attempt to apply an array to the triangular prism problem indicated that her understanding of arrays was specific to the rectangular prism. We had asked her if the "sheets" idea applied to the triangular prism, and she had trouble imagining cutting a unit cube to fit an acute angle:

> **Amelia:** Where it's a triangle, you obviously can't squeeze a square into an acute angle. I guess that's why we have formulas, because we can't physically put a cubed object into that space there.

A thorough understanding of arrays would include the idea of fractions of unit cubes, an idea that eluded Amelia. She had found the correct volume of the rectangular prism, but found surface area in the triangular prism task. Geddy's work was similar: she explained volume as an array for the rectangular prism but found surface area for the triangular prism (see Figure 5). In contrast to Amelia, however, Geddy did seem to apply the array model to the rectangular prism. She said

> **Geddy:** Well, since it's 108, it's an even number of cubes. You'd be able to use squares equal to volume 1 ft cubed and you should be able to fit them all in without having any gaps.

We take Geddy's phrase about "squares equal to volume 1 ft cubed" to mean unit cubes and the statement about "fitting them in without gaps" to be indicative of an array of cubes.

A second finding about calculus students is that some do not have an array model for volume at all. Carly, one of the students who found the surface area of the rectangular prism, did not seem to have an understanding of arrays. She discussed her reasoning about the rectangular prism:

> **Carly:** I know there's an equation for volume. I don't remember what the equation is, but you know this is 5, this is 4, this is 10 [labels diagram]. I know you can find each side but I don't think that gives you the volume… like find 10 times 4 so you know this side is 40 cm and this side is 10 times 5 so this side is 50. And I would just assume that this side is the same so I'd say the back side is 50 and the bottom would be 5 times 4 so that would be 20 and the top would be 20. But if you add all those together I don't think that would give you the volume because volume includes all the space in between – like in the middle of the box.

Carly had the correct idea that volume should account for "the middle" of the shape, but was not able to extend the idea of accounting for "the middle" to appropriate mathematics. Rather, she reverted to two-dimensional ideas. This leads us to the following: we hypothesize that there may be a relationship between students' array understanding and their success on these tasks.

This finding is strengthened by a number of students who described volume as an array or as layers and answered all of the problems correctly. For instance, Luke found the correct volumes for all shapes. He talked about depth and planes in the rectangular prism, which we consider analogous to layers in an array:

> **Luke:** The volume of the box is 10 cm * 4cm * 5 cm, and that is 200 cm^3. I think of it as having an area, which is one plane, and you're multiplying it over 4 cm so you multiply your one plane by the depth of the object and that gives you the volume.

Wendell, who also found volume on all the tasks, discussed the volume of the rectangular prism similarly:

> **Wendell:** I'm a hands-on kind of person so I think it would be easiest to explain by giving them 5 one-centimeter cubes and show them that's one stack, then do it by 4, then tell them there are 10 stacks high. Then you tell them if there are 20 in the

bottom and 10 stacks high, 20 * 10 is how you find the volume.

We did not ask Luke and Wendell to sketch arrays for the other shapes, but based on their descriptions for the rectangular prism array and their success on the other task, we suspect they would have sketched and described accurate array models for these shapes. Furthermore, we suspect that having these models is related to their computational success across all of the tasks. In contrast, the less-robust understanding of arrays in the other students (Geddy, Amelia, and Carly) may be related to their surface area finding in other tasks. Carly found surface area on all the tasks and did not understand arrays; Geddy found surface area on two of the tasks but did understand arrays; Amelia understood an array only in terms of the rectangular prism and found volume for that task but surface area on the others. Luke and Wendell understood arrays and solved all of the problems correctly. We believe these data suggest that having an array model of volume for a shape has some connection to successfully finding volume of that shape (but not necessarily others), while not having an array model of volume for a shape may be connected to surface area computations (or computations involving both surface area and volume elements).

These calculus students' difficulties with arrays are similar to the difficulties faced by elementary students about the same topic. Elementary school students have trouble understanding the unit structure of an array (Battista & Clements, 1996) and using an array to compute volume (Curry & Outhred, 2005). We found that some calculus students, like elementary school students, have trouble with the structure of an array (e.g., Carly's work). However, while elementary school students often do not see the relationship among the rows, columns, and layers in an array, many calculus students do (e.g., Wendell and Luke) and can use them for computation. Between the extremes of 'no array model' and 'array model', there are calculus students who have a model for a rectangular prism but not other shapes. We suspect a student's array model, robust or otherwise, is related to computational ability. In conclusion, while some calculus students' have overcome the difficulties elementary school students face with arrays, others continue to struggle with array models and their use in volume computation.

In the next section, we state some final conclusions as well as implications for instruction and suggestions for further research.

Conclusions, Implications for Instruction, and Suggestions for Further Research

One of the questions this study sought to answer was "How successful are calculus students at solving computational volume problems?" Success depends on shape. For the rectangular prism, 98.5% of calculus students found volume; 94.5% of students found the volume of the cylinder; and 84.2% of students found the volume of the triangular prism. It's important to note that students were more successful with the rectangular prism than with the assumedly less familiar cylinder and triangular prism.

This may have implications for volume-finding in calculus; for instance, volumes of solids of revolution are rarely elementary shapes. A related finding with relevance to volumes of solids of revolution is that students who successfully found volume often thought of it in terms of *area of base times height* or as an array with layers. The base or a layer is a cross-section of the solid. Solving a volume of revolution problem often requires identifying the shape of a cross section and an expression to represent its area. The area expression is then integrated to find the volume of the solid. We think this is similar to *area of base times height* and to a layer model of volume because the integration sums the volume of infinitessimally thin layers (cross-sections). We suggest that instructors include these models of volume as part of instruction about volumes of solids of revolution.

The other research questions concerned whether or not calculus students find surface area when directed to find volume. As with volume-finding, the percentage differs by shape (1.5% for the rectangular prism, 5.5% for the cylinder, and 15.2% for the triangular prism). Further research is needed to know exactly what causes the differences in surface area finding for different shapes, but findings from the present study provides some insights as to why students find surface area at all. One reason is that some students think that adding the areas of faces measures three-dimensional space. A second is that some students have a clear conceptual understanding of volume, but blend of surface area and volume elements in computational formula. This has important implications for calculus learning, particularly in optimization problems. Standard optimization problems require students to minimize the surface area for a given volume (or vice versa). Students must construct formulae for both, solve one for a variable (often height) that can be substituted into the other equation, and only *then* can a student begin the calculus involved. It's possible that difficulties with optimization may be linked to these first few non-calculus steps. While further research is needed to confirm if this is the case, we suggest that instructors provide opportunities for students to revisit surface area and volume concepts and formulae, and perhaps give students these formulae on exams to ensure that calculus knowledge, rather than geometry knowledge, is tested.

Issues shared by calculus students and elementary school students

Finding surface area when directed to find volume is an issue that has been documented with elementary school students. In a Battista and Clements (1998) study, three tasks were given in which third- and fifth- grade students were directed to find the volume of a three-dimensional array of cubes. About 18% found surface area using pictures of a 4x2x2 array, a 4x3x3 array, and a manipulative 3x4x5 array. In this study, 1.5% of students found surface area of a picture of a rectangular prism. We conclude that, at least for this shape, calculus students are more successful than elementary school students at finding volume, but it exists as an issue in both populations.

An additional issue shared by calculus students and elementary school students is that some students from both populations struggle with representing volumes with arrays and using the arrays for volume computations.

Implications for instruction

In addition to the instructional implications mentioned above, we think instructors can use students' computational formulae to diagnose their ideas. Our findings indicate that students' formulae are often indicative of ideas they hold about surface area and volume, such as Nell and Jo's thoughts about whether or not to include the two bases in finding the area of the cylinder. We think that sorting students' formulae, as we did in Table 3, can be useful to identify ideas they bring to a computation. Instruction can then target those particular conceptions.

Additionally, we believe that many of students' errors result from not understanding volume as an array. We thus suggest that instructors provide students with educational opportunities to model volumes with arrays and connect the models to volume formulae. In a smiliar vein, we think that the conception of volume as *area of base times height* should be emphasized and, in calculus, connected to the idea of cross-sections. This could improve student success on volumes of solids of revolution, a notably difficult calculus topic (Orton, 1983).

Finally, we found that students' success in finding volume was somewhat shape-dependent. We suggest that volume learners practice finding the volumes of a variety of shapes, rectangular solids and otherwise.

Suggestions for further research

We are particularly interested in how students' understanding of volume, and the surface area-volume combinations found here, are brought to bear in calculus topics like optimization, related rates, and volumes of solids of revolution. Many optimization problems use both surface area and volume, and we are interested in how students who have difficulty with volume work through these problems. We suspect that, as in other areas of research about calculus learning, the issues students have with calculus topics is rooted in issues with underlying concepts. Further research is needed to investigate if this is also the case with volume and the calculus topics that use it.

Works Cited

Battista, M.T., & Clements, D.H. (1998). Students' understandings of three-dimensional cube arrays: Findings from a research and curriculum development project. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 227-248). Mahwah, NJ: Erlbaum.

Battista, M.T., & Clements, D.H. (1996). Students' understanding of three-dimensional rectangular arrays of cubes. *Journal for Research in Mathematics Education, 27(3),* 258-292.

Byrnes, J.B.. (2000). Cognitive Development and Learning in Instructional Contexts (2nd Edition). Pearson Allyn & Bacon.

Carlsen, M. (1998). A cross-sectional investigation of the development of the function concept. In Schoenfeld, A., Kaput, J. & E. Dubinsky (Eds.), Research in Collegiate Mathematics Education.

Curry, M. & Mitchelmore, M. (2006). Development of children's learning of length, area, and volume measurement principles. Mathematics in the centre: Proceedings of the 30th conference of the International Group for the Psychology of Mathematics Education Charles University, Faculty of Education, Prague, Czech Republic 16-21 July 2006, Vol. 2, p.377-384.

Curry, M., & Outhred, L. (2005). Conceptual understanding of spatial measurement. In P. Clarkson et al. (Eds.), Building Connections: Theory, Research and Practice (Proceedings of the 27th MERGA conference), Sydney: MERGA.

Davis, E. (1973). A study of the ability of selected school pupils to perceive the plane sections of selected solid figures. *Journal for Research in Mathematics Education, 4(3),* 132-140.

De Corte, E., Verschaffel, L., & Van Collie, V. (1988). Influence of number size, problem structure, and response mode on children's solutions of multiplication word problems. *Journal of Mathematical Behavior; 7*, 197-216.

Dorko, A. (2011). Calculus students' understanding of area and volume in non-calculus contexts. Unpublished masters thesis, University of Maine at Orono.

Fuys, D., Geddes, D., & Tischler, R. (1988). The van Hiele model of thinking in geometry among adolescents. *Journal for Research in Mathematics Education Monograph* 3. Reston: NCTM 1988.

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory*. Chicago: Aldine.

Hirstein, J., Lamb, C., & Osborne, A. (1978). Student misconceptions about area measure. *The Arithmetic Teacher, 25*(6), 10-16.

Hunting, R. (1997). Clinical interview methods in mathematics education research and practice. *Journal of Mathematical Behavior, 16(2),* 145-165.

Izsák, A. (2005). You have to count the squares: Applying knowledge in pieces to learning rectangular area. *Journal of Learning Sciences, 14,* 361-403.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G. Martin, &

D. E. Schifter (Eds.), A Research Companion to Principles and Standards for School Mathematics. Reston, VA: National Council of Teachers of Mathematics.
Lehrer, R, Jenkins, M., & Osana, H. (1998). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 137-167). Mahwah, NJ: Erlbaum.
Mack, (2001). Building on informal knowledge through instruction in a complex content domain: Partitioning, units, and understanding multiplication of fractions. *Journal for Research in Mathematics Education, Vol. 32(3),* 267-295.
Martin, T. (2000). Calculus students' ability to solve related rates problems. *Mathematics Education Research Journal*, 12(2) p. 74-91.
Milovanović, M., Takači, Đ., & Milajić, A.(2011). Multimedia approach in teaching mathematics - example of lesson about the definite integral application for determining an area. *International Journal of Mathematical Education in Science and Technology*, 42: 2, 175-187, First published on: 17 January 2011 (iFirst).
Nesher, P. (1992). Solving multiplication word problems. In G. Leinhardt, R. Putnam, & R.A. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 189-219). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics, 14,* 1-18.
Peled, I., & Nesher, P. (1988). What children tell us about multiplication word problems. *Journal of Mathematical Behavior, 7* , 239-262.
Rasslan, S., & Tall, D. (2002). Definitions and images for the definite integral concept. in: A. D. Cockburn & E. Nardi (eds.) *Proceedings of the 26th Conference PME*, Norwich, 4, 89-96.
Redish, E.F. (1997). Student difficulties with math in the context of physics. Available online at: http://www.physics.umd.edu/perg/papers/redish/talks/math
Simon, M., & Blume, G. (1994). Building and understanding multiplicative relationships: A study of prospective elementary school teachers. *Journal for Research in Mathematics Education, 25,* 472-494.
Thompson, P., & Silverman, J. (2006). The concept of accumulation in calculus. . In M. Carlson & C. Rasmussen (Eds.). *Making the connection: Research and teaching in undergraduate mathematics education.* (pp.27-41). MAA Notes #73. Mathematical Association of America: 2008.
Tomlinson, K. (2008). The impact of cooperative guided reflection on student learning: the case of optimization problem solving in Calculus 1. Disponible. 30 April 2010 http://minds.wisconsin.edu/handle/1793/24574.
Trigueros, M. & Ursini, S. (2003). First year undergraduates' difficulties in working with different uses of variable. *CBMS Issues in Mathematics Education 12*, 1-29.
U.S. Department of Education (2007). NAEP Exam Questions and Statistics. Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Mathematics Assessment.
White, P., & Mitchelmore, M. (1996). Conceptual knowledge in introductory calculus. *Journal for Research in Mathematics Education, 27*(1) 79-95.

PROOF STRUCTURE IN THE CONTEXT OF INQUIRY BASED LEARNING

Alyssa Eubank
Shawn Garrity
Todd Grundmeier
California Polytechnic State University, San Luis Obispo

An introductory proofs course was taught using Inquiry-based learning (IBL), which gives authority to students and allows them to present to their peers rather than having the instructor as the focus of the class. Data was collected from the final exams of 68 students and analyzed based on proof structure. Problems included concepts that were introduced prior to and during the course. This research utilizes an adaptation of Toulmin's method for argumentation analysis. Our goal was to compare the proof structures from these students to previous research that applied Toulmin's layout to mathematical proof. There was a much wider variety of proof structures than expected, which could be a result of the IBL atmosphere.

Key words: Proof, Toulmin, Inquiry-based learning, Structure

Introduction

Understanding student thought processes has been given much attention in mathematics education research. Furthermore, the way in which a student constructs a proof is a key indicator of how they organize their thoughts to approach problems. This paper discusses a tool for analyzing the structure of undergraduate proof and explores characteristics of proofs in the context of an introductory proof course.

An adaptation of the inquiry based learning method was used to teach three sections of Methods of Proof, an introductory mathematical proofs course. Students presented problems to the class and critiqued their peers. At the end of the quarter, students were required to complete a comprehensive final exam. Our data consists of problems from the final exams of 68 students, which were each graded by the instructor.

Inquiry Based Learning

Currently, a typical undergraduate mathematics classroom consists of instructor given lectures with little to no student presentation. Inquiry based learning (IBL) instead places responsibility on students by encouraging students to discover mathematics for themselves rather than simply relying on the authority of an instructor. In this setting peer collaboration, peer critique, and student presentations are common.

Toulmin Overview

Stephen Toulmin analyzed argumentation and developed a new approach to formal logic. He classified six interrelated components he believed essential in constructing a sound and convincing argument: claim, grounds (data), warrants, backing, rebuttal, and modal qualifiers (Toulmin, Rieke, and Janik, 1979).

In Toulmin's (1979) work, an argument was defined as "the sequence of interlinked claims and reasons that, between them, establish the content and force of the position for which a particular speaker is arguing" (p.13). This definition differs from that of mathematical proof. Argumentation relies on making claims and then justifying, while mathematical proof relies on making inferences of previous results to come to a claim (Barrier, Mathe, & Durand-Guerrier, 2009). In other words, argumentation relies on the content of each claim and proof relies on the function of each claim. (In the context of this paper, we use mathematical argument and proof to mean the same thing). Though Toulmin did not devise his scheme from the perspective of mathematical proof, it is valuable in this context because of the many parallels between argumentation and proof.

Merits of Toulmin

Many math education researchers agree that a formal proof contains elements identified by Toulmin, as evidenced by the various proof analysis schemes adapted from Toulmin's work. Though much merit is given to Toulmin's scheme, it is often the case that a *restricted* version is applied in the context of proof. In this scheme, the rebuttal and qualifier are not considered to be a part of mathematical argumentation. However, Inglis, Mejia-Ramos, and Simpson (2007) argue that you need Toulmin's *complete* model to analyze proof (i.e. consider all six elements). In their research, an interviewer interacted with a student in order to understand the process by which they came to their conclusion. This interaction enabled researchers to witness corrections of mistakes and possible uncertainty of each student (i.e. rebuttals and qualifiers of their arguments). In this case, the *complete* scheme can be considered necessary. On the contrary, a researcher analyzing written proof only witnesses the final stage of the student's thought process; students have effectively already qualified their statements and considered plausible rebuttals and do not present any uncertainty. Here, since qualifiers are no longer being stated, a *restricted* scheme is sufficient.

In this paper we propose that, by Toulmin's (1979) definition, a *qualifier* is not directly associated with proof and hence suggest a new definition to use within the context of proof. According to Toulmin, *qualifiers* determine the strength of an argument in that they restrict the situation in which the final claim is true. In a mathematical proof, the final claim should be true in every situation, but its validity can rely on sub-cases within the proof. Hence, the definition of *qualifier* for this research deals with cases and will be discussed more in depth later.

Previous Results

Toulmin's criteria for the structure of argumentation have been used in various contexts of mathematics education research: traditional lecture-based classrooms (Fukawa-Connely, 2011), interviews (Inglis, Mejia-Ramos, & Simpson, 2007), and classroom discussions (Krummheuer, 2007). Toulmin's scheme has also been applied to everyday argumentation, such as that found in the workplace (Simosi, 2003). One common finding in this line of research is a lack of warrants and backing within an argument or proof. Fukawa-Connely (2011) analyzed student proofs collected from a lecture-based abstract algebra course. Probably due to the rigorous level of the course and the students' experience with proof, it was observed that the instructor often left warrants out of her written proofs and rarely wrote backing. In one particular proof, Fukawa-Connely (2011) found a major lack of backing where only 25% of the claims were supported by warrant. This seemed to have been copied in the students' work. However, Fukawa-Connely's (2011) further research led him to conclude that there exists a correlation between unfamiliarity of content and the amount of detail (i.e. warrants and backing) within a proof. As topics became more unfamiliar, it appeared that students felt more warrants were necessary to complete their proof.

A possible reason for lack of backing was discussed by Simosi (2003) pertaining to argumentation in the business workplace. It was observed that 48 out of 66 arguments contained some sort of stated warrant or backing. When the arguer did not explicitly state the backing, it was often because s/he deemed the information obvious and hence not necessary to conclude; Simosi (2003) referred to this as *self-evident backing*.

Researchers also found a difficulty in distinguishing between warrants and backing. In the specific case where Toulmin's scheme was applied in the workplace, the language used within an argument created this confusion. The wording of a statement was sometimes misleading (i.e. the arguer used wording typically used to introduce a warrant, but the statement itself was functioning as backing) (Simosi, 2003).

While the structure of mathematical proof and argumentation has been explored in varied contexts, little or no research exists about the structure of student proof in the context of an IBL mathematics course. This work attempts to fill this gap in the literature.

Methodology

Previous Definitions

Toulmin's statement classifications have been applied to mathematical argument by multiple researchers. Each time, the definitions varied slightly depending on the context, though the general meanings were relatively similar. The following is a summary of the various ways each classification has been defined:

<u>Data</u>: These are undoubted statements (Krummheuer, 2007). They are also the foundation upon which the argument is based, often identified by directly following the words "if" or "let" (Fukawa-Connely, 2011).

<u>Warrant</u>: These are statements that contribute to the legitimizing (Krummheuer, 2007), meaning they justify the relationship between the data and conclusion by drawing upon previously demonstrated facts (Fukawa-Connely, 2011). Warrants are often algorithmic, stating the procedure that led to the claim (Rasmussen & Stephan, 2008).

<u>Backing</u>: This classification is for statements that refer to the permissibility of the warrant (Krummheuer, 2007), suggest why it is valid, and explain the applicability of the warrant (Fukawa-Connely, 2011). They can also be thought of as answers to the question, "Why should I accept your argument as mathematically sound?" (Rasmussen & Stephan, 2008).

<u>Qualifier</u>: These are expressions of the degree of confidence of the conclusion, but they also suggest that generalizability of the claim may be limited (Fukawa-Connely, 2011).

<u>Rebuttal</u>: Rebuttals go along with qualifiers, in that they state conditions under which the conclusion would not hold (Fukawa-Connely, 2011).

<u>Conclusion/Claim</u>: These are the statements that are being argued, often following the word "then", or presented as a deduction from data (Fukawa-Connely, 2011).

Definitions

After careful consideration of Toulmin's (1979) definitions and those of past research on mathematical proof, we agreed on the following definitions to classify statements in student proof.

Statements classified as *data* are the undisputed facts that begin an argument, usually the information given in the statement of the problem (e.g. "let f be a function"). *Claims* are the series of steps contained in a proof used to reach the final conclusion (e.g. "then f is onto").

Warrants are those statements that justify a relationship, which would often be algebraic steps or an indication that a definition was used (e.g. "by definition of function"). The *backing* for each warrant is a discussion of why the warrant is permissible. Often this would be restating previously accepted information in a different way, or following the word "recall". The question "Why is this allowed?" was used to determine whether a statement should be identified as *backing* by searching for a description of how the warrant applies (e.g. a backing could be a restatement that the qualifications of a 1-1 function are met, which would support a warrant using the definition of 1-1).

It can be confusing when trying to distinguish between warrants and backing, since both are used as support. Others have had the same difficulty, which is one reason why critics claim this is an unrealistic scheme for real-life arguments, noting that it was sometimes unclear whether the arguer was giving a warrant or backing to the claim (Simosi, 2003). In order to distinguish correctly, one must focus on the function of the statement. Warrants are the tools used and backings are the reasons those specific tools work. Consider the problem of having to attach one piece of wood to another.

When it comes to *qualifiers* and *rebuttals*, the original use does not apply to this research, since students do not intentionally express uncertainty in their written statements on a final exam. However, *qualifiers* can also express the limitations of a statement's generalizability. Hence, we chose to identify each sub-case within a proof as a *qualifier* because the arguments that follow are limited to that specific case. *Rebuttals* were not used at all in our coding scheme.

Coding Scheme

The proofs of the statements in the coded problems required several steps. Thus proofs consisted of a string of claims, each with its own warrants and backing. For every student's proof, the arguments were mapped using a similar schematic to one often seen in research of this nature.

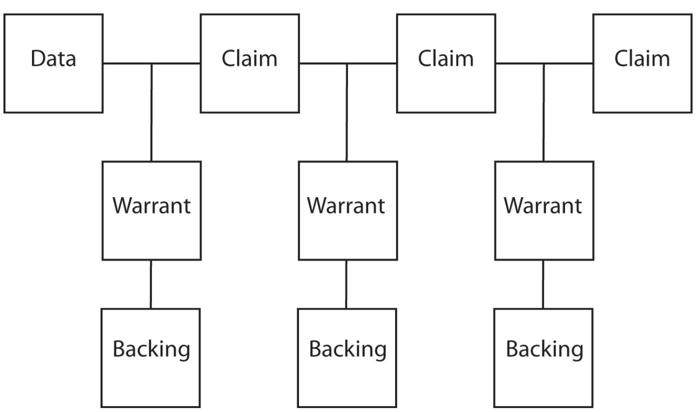


Figure 1: General schematic of coding scheme

After mapping each proof, the following characteristics were recorded:

Length: this refers not to the number of words used in each proof, but rather to the number of steps taken. Thus the total number of claims was counted. When recording length, a designation of *short* (S – one or two claims), *average* (A – three or four claims), or *long* (L – five or more claims) was assigned to each proof. Note that the word "average" does not refer to the mathematical average, but rather to a typical length. If a student did not write any claims, the proof was called "other."

Existence of warrants: this indicates how well the proof was warranted. Each proof was identified as *complete* (*c* – warrants given for 100% of the claims), *most* (*m* – more than 50%, but not all, of the claims are warranted), *limited* (*l* – at most 50% of the claims are warranted), or *none* (*n* – 0% of the claims are warranted).

Existence of backing: this indicates how well the proof was backed. Each proof was given a designation of *complete* (*c* –backings supplied for every warrant), *limited* (*l*– some backings provided, but not for every warrant), or *none* (*n* – no attempted backing).

Qualifiers: these are sub-cases within a proof. The specific number of sub-cases was not recorded. Since the use of qualifiers is in essence a claim that sub-cases are appropriate, all qualifiers together count as one claim for the purpose of coding the length of a proof.

Each proof was given a code in every category, resulting in a three-letter designation. For example, a proof with *average* length, *complete* warrants, and *limited* backing would be given a designation of Acl. *Qualifiers* and incorrect statements or implications were tallied for each proof as well. In addition, the instructor graded the finals at the end of the course, giving each problem a score from zero to ten. These scores were included in the analysis because they were a good indication of the acceptability of the proof in the eyes of the instructor.

Coded Problems

For the purposes of this research, we chose problems from the final exam to evaluate using the above coding scheme related to the structure of mathematical proof. When looking at the various concepts covered throughout the course, we chose to analyze an integer problem and a function composition problem. The integer problem represents a familiar concept that students had seen prior to the class, while the function problem represents a concept that was new to students.

Two versions of the final exam were administered to students, both written by the instructor of the course. Although the two versions contained different problems, they were identically structured and were comparable in difficulty. On each version, students were asked to prove propositions involving properties of integers and functions, among other concepts. Two problems were chosen from each final exam, one from the Integer category and one from the Function category. The number of students given each problem is noted in brackets.

Integer problems:

Let x be an integer. If 8 does not divide $x^2 - 1$, then x is even. [48]
Let x and y be integers. If xy is even, then x is even or y is even. [20]

Function problems:

Let $f : A \to B$ and $g : B \to C$. If $g \circ f : A \to C$ is 1-1 and f is onto, then g is 1-1. [48]
Let $f : A \to B$ and $g : B \to C$. If $g \circ f : A \to C$ is onto and g is 1-1, then f is onto. [20]

Results

Structure Codes

Since the coding scheme constructed for this project was based on proof length, warrants, and backings, there were 28 potential codes. Twenty of these arose in the coding of this particular data set of proofs. The intention of this section is to provide an overview of where these codes appeared.

When looking across the five most frequently used codes for Integer and Function problems, there were three common codes: Acl, Aml, and Lml. Refer to Table 1.

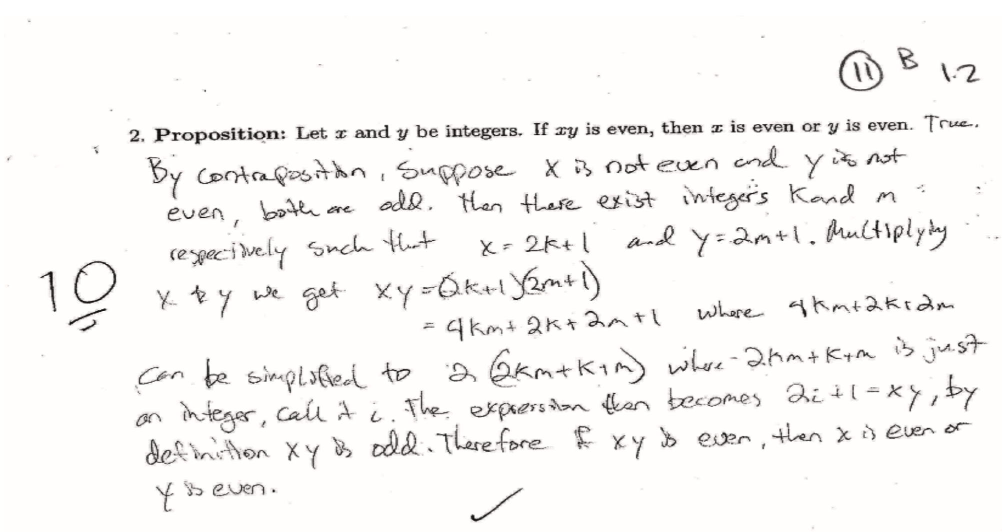
Table 1: Most common structure codes

Integer Problem						
Structure Code	Acl	Acc	Amc	Aml	Lml	
Frequency	21	10	9	9	5	

Function Problem						
Structure Code	Scc	Acl	Lml	Lcl	Aml	Lcc
Frequency	11	9	7	6	5	5

All Problems						
Structure Code	Acl	Aml	Scc	Acc	Lml	
Frequency	30	14	13	12	12	

The most common code was Acl. This was often a proof that was correct, but the very last step was not given any backing. To prove the integer proposition, many students chose to use contraposition. For the final claim that the proposition is true, many of the students whose proofs were coded as Acl stated that they used contraposition as their warrant but did not give backing. In this case, the backing would have been writing out that the contrapositive is true. Figure 2 below shows a student's proof for which this is the case, followed by the proof's schematic.



11 B 1.2

2. Proposition: Let x and y be integers. If xy is even, then x is even or y is even. True.

By contraposition, Suppose x is not even and y is not even, both are odd. then there exist integers k and m respectively such that x = 2k+1 and y = 2m+1. Multiplying x & y we get xy = (2k+1)(2m+1)
= 4km+2k+2m+1 where 4km+2k+2m can be simplified to 2(2km+k+m) where 2km+k+m is just an integer, call it i. The expression then becomes 2i+1 = xy, by definition xy is odd. Therefore if xy is even, then x is even or y is even.

10

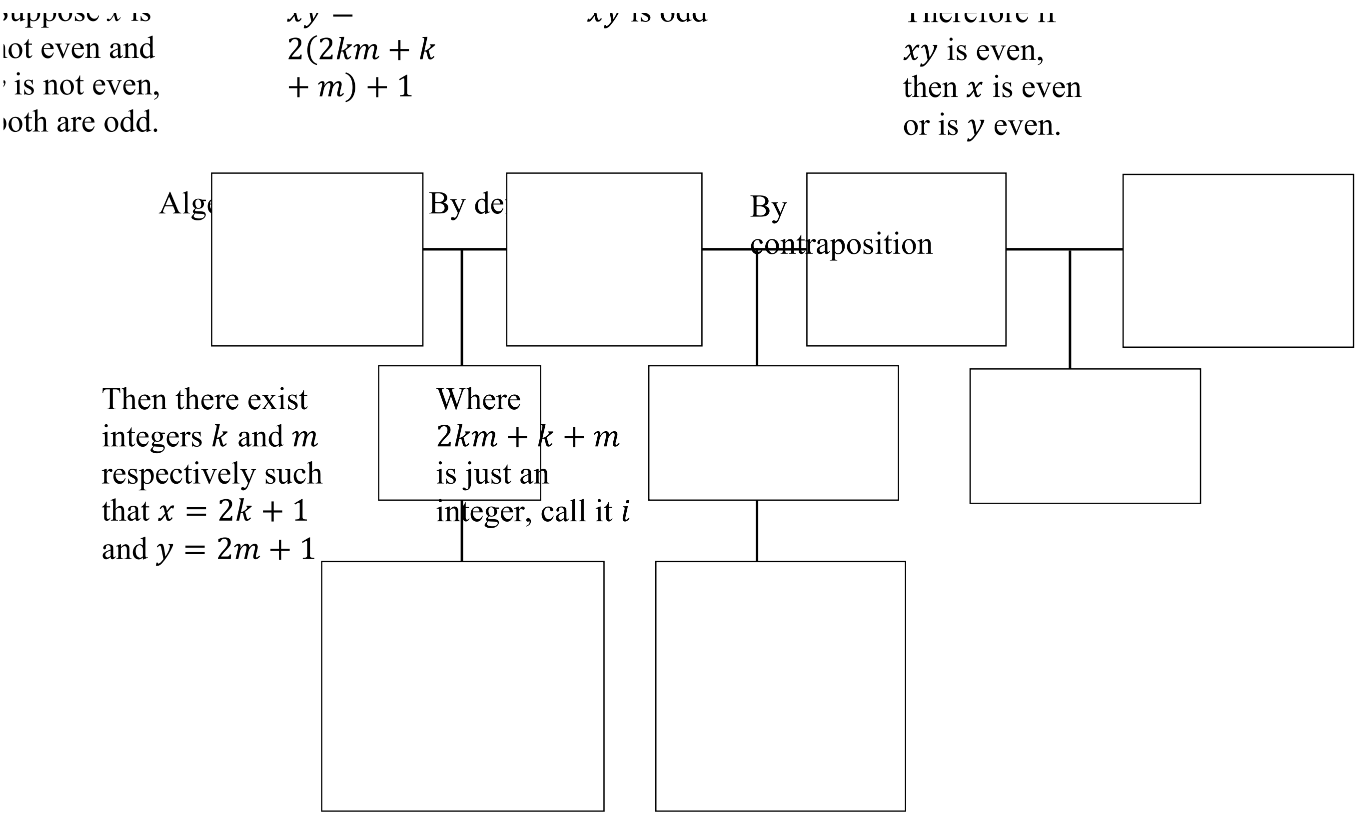


Figure 2: An Acl proof from the Integer category

Figure 3 illustrates the way a student with a similar proof backed their final claim, thus resulting in an Acc code.

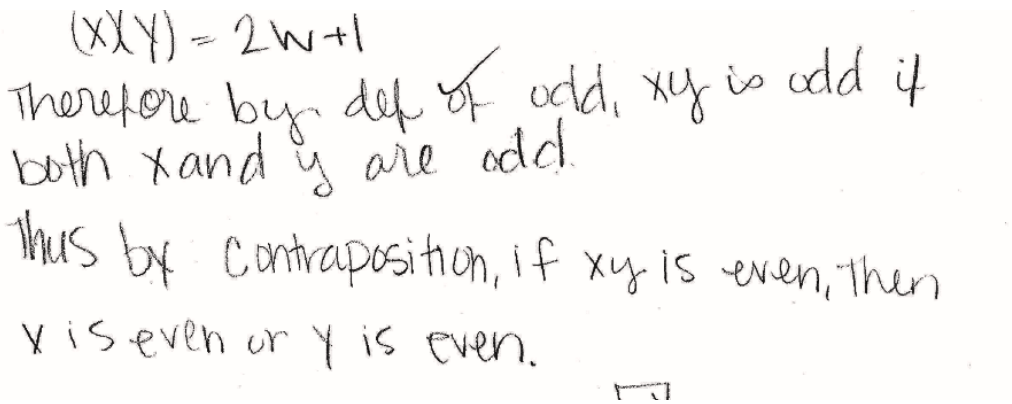


Figure 3: Conclusion of an Acc proof from the Integer category

The codes Acc, Lcc, and Lcl were among the five codes with the highest average scores in both the Integer and the Function category. In the entire data set, two of these five codes have both *complete* warrants and *complete* backings. Table 2 illustrates this result.

Table 2: Structure codes with highest average scores

Integer Problem					
Structure Code	Lcc	Acc	Lmc	Lml	Scl
Average Score	10.00	9.40	9.25	9.00	9.00

Function Problem					
Structure Code	Lcc	Lcl	All	Lmc	Acc
Average Score	10.00	9.33	9.00	9.00	7.50

All Problems					
Structure Code	Lcc	Lmc	Acc	Scl	All
Average Score	10.00	9.14	9.08	9.00	8.75

Of all the Length codes identified, there was a slightly greater variety of codes with *average* length (8 codes) than of those with *long* (6) or *short* (5) length. When observing the various Warrant codes used, there were more codes containing *complete* (7 codes) than *most* (6), *limited* (6), and *none* (0). Finally, the same number (7) of *complete* and *limited* Backing codes was observed in the data. It is interesting to note that all possible codes containing *most* (in the Warrant category) were represented by some student proof.

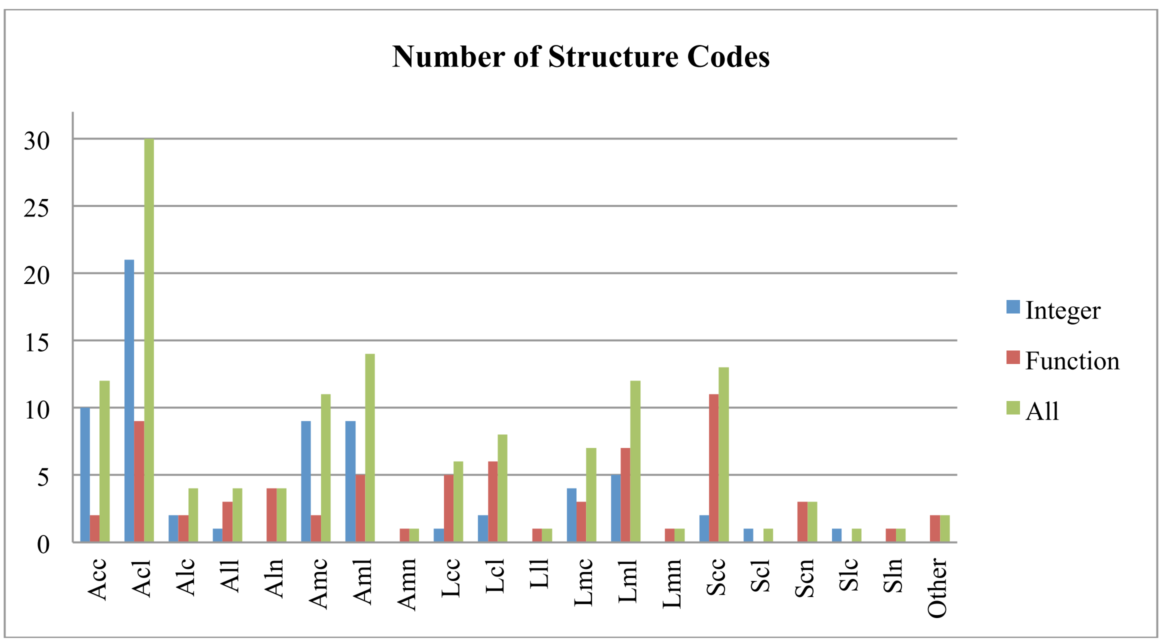


Figure 4: Number of each structure code by Integer and Function categories

When looking at the total number of each structure code observed in the data (Figure 4), the Integer proofs appear to have a more consistent structure than the Function proofs. This can be seen in the distribution of the number of each code. The Integer proofs were associated with 13 structure codes while the Function proofs were associated with 18. The code Acl accounted for more than 30% of the Integer proofs. The next most frequent code (Acc) accounted for slightly less than 15%, about half as often. The appearance of the two most common codes (Scc appearing 16% and Acl appearing 13% of the time) for Function proofs only differs by 3%, which once again highlights in inconsistency among the Function proof structures.

Warrants

Providing warrants for claims is one of the most important parts of a proof. It was thus encouraging to see that every coded proof had some kind of warrant (other than the two students' proofs which had no claims to warrant in the first place). Recall that each proof was coded depending on the percentage of the claims that had a supporting warrant. If every claim was warranted, the proof was given a designation of *complete* for warrants. If more than half of the claims were warranted, the proof was given a *most*. Proofs coded as *limited* had half or less of the claims supported with a warrant. There was a fourth code to designate warrants, which was *none*, meaning that none of the claims in a proof were warranted. However, this code was not used, since every proof had at least one of its claims supported.

Even further, in both Integer and Function categories, the majority of the proofs were *completely* warranted. While the numbers of proofs with *complete* (c) warrants were similar between the two categories (37 and 36, respectively), there were large differences in the distributions of the *most* (m) and *limited* (l) proofs. Figure 5 offers a comparison of these codes by Integer and Function categories.

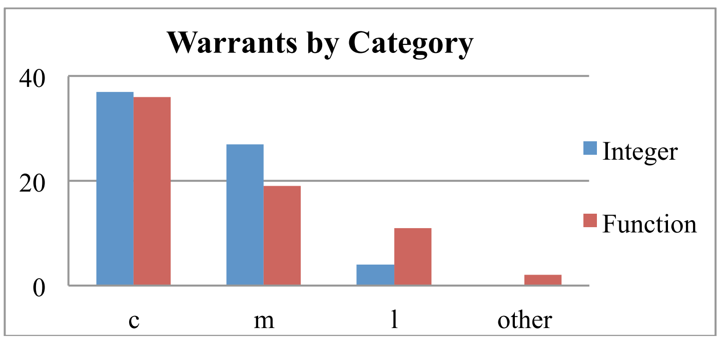


Figure 5: Number of proofs with c, m, and l Warrant codes

The Integer category has 23 more *m* proofs than *l* proofs, whereas in the Function category there is a difference of only 8. Also note that the two "other" proofs (which had no claims and thus could not be coded based on number of claims warranted) were both in the Function category.

As expected, the average score of proofs with *complete* warrants was higher than that of the other proofs. Table 3 below displays this information.

Table 3: Average score for Warrant codes in each category

Average Score by Warrant Code			
	c	m	L
Integer	8.49	8.04	8.00
Function	6.58	6.21	6.45
All Problems	7.55	7.28	6.87

There does not appear to be any correlation between score and proportion of warrants, however. Thus while a completely warranted proof is more likely to have a higher score, the converse is not necessarily true. It may be interesting to note, however, that the two most common scores (10 and 3) had very similar distributions of warrant codes. See Figure 6 for an illustration.

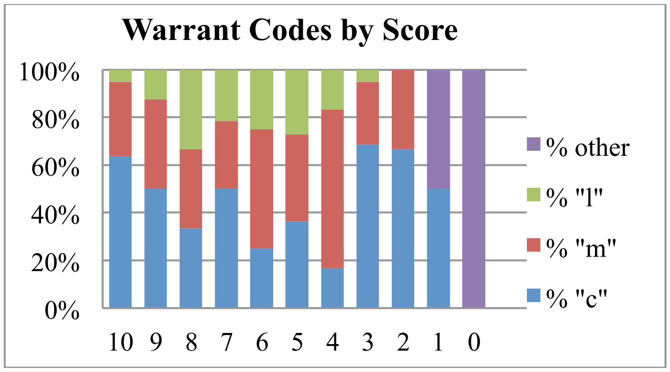


Figure 6: Distribution of Warrant codes for each score

Backing

As with warrants, all Integer proofs had some kind of backing, which gives validation to the warrants. The Function proofs, however, had 10 proofs with no backing present at all, not including the two "other" proofs.

In contrast to the results found concerning warrants, *complete* backing was not the most common occurrence. In both Integer and Function proofs there were more *limited* proofs than *complete* ones, as shown in Figure 7.

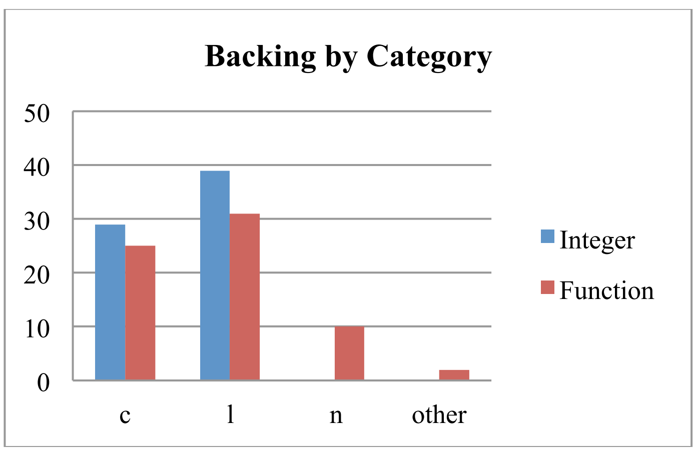


Figure 7: Number of each category with c, l, n, or other Backing codes

This most common code (*l*) also had the highest average score overall (7.64). The *completely* backed proofs, on the other hand, had a slightly lower average score (7.43), which was somewhat unexpected. However, the proofs with no backing still had the lowest average score (5.3), which was expected. Table 4 displays these results.

Table 4: Average score for Backing codes in each category

Average Score by Backing Code			
	c	l	n
Integer	8.45	8.15	—
Function	6.24	7.00	5.30
All Problems	7.43	7.64	5.30

Discussion

Structure Codes

A large part of this IBL classroom was student presentation and peer collaboration. It could be expected that this level of collaboration would influence students to approach and prove statements with similar structure, but the data from this project seems to provide evidence to the contrary. It was found that there was not a consistent proof structure across all analyzed proofs. The most common structure code appeared only 30 times (22.1% of all proofs) and the next most common code only appeared 14 times (10.3% of all proofs). This lack of consistency may be due to the fact that the class did not have one consistent person modeling formal mathematical proof but instead an entire classroom of presenters.

In Fukawa-Connely's (2011) work, it was found that in a traditional, lecture-based classroom, students modeled their proofs after an authoritative figure (i.e. the instructor). It may be the case that the IBL students did not model their proofs after their figures of authority because they did not fully trust each student presenter's mathematical competency. This lack of trust would likely force students to critically think and assess the validity of each proof instead of trusting that the professor is correct. Also, as students are being exposed to the many different proof structures provided by various presenters, they are able to judge which proof structure makes the most sense and works best for them. The wide variety of proof structures identified in this data set suggests that an IBL classroom facilitates a flexible environment that encourages student creativity within formal proof. The fact that average scores were not drastically different between structures also supports this.

Further inconsistency of proof structure was found when comparing the average scores of structure frequency (Figure 8), where the frequency of a structure code is defined as the number of times a particular code appeared in the data set. Though there appears to be a positive correlation between the average score and frequency, such a small correlation coefficient (0.07) implies that a significant correlation does not exist. Once again, this could possibly be explained by students' response to having many different models and authority figures.

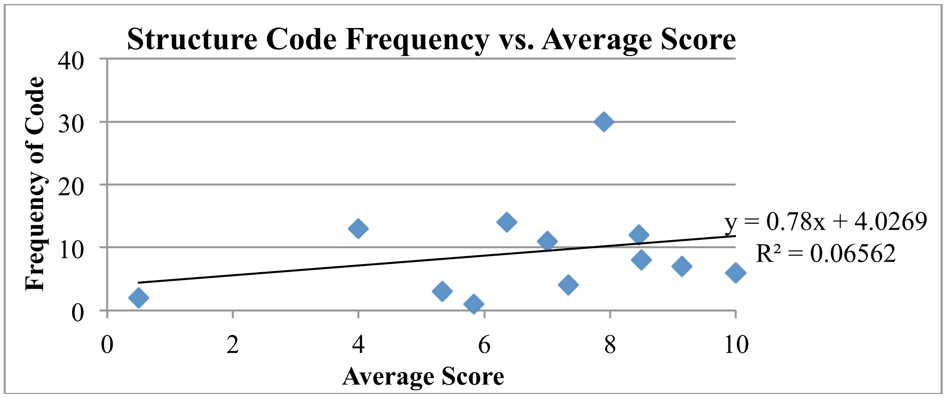


Figure 8: Average scores of proofs based on the number of times each structure code appeared

Warrants

The majority (almost 54%) of proofs were completely warranted, both from problems that were more familiar to the students (Integer) and from those that were unfamiliar (Function). This is evidence that the students in this IBL class learned the importance of providing explanation for their claims, not simply stating them. Recall that Fukawa-Connely (2011) found that only about 25% of the claims of a particular proof in the class he observed were warranted. This is not directly comparable to the IBL class from this research since the two are different courses and represent different levels of mathematical maturity. But Fukawa-Connely's (2011) research still gives a reference point, leading to the conclusion that the students from this class successfully learned to warrant their claims.

However, even though students learned to provide warrants, the warrants they provided were not always sufficient. A statement was counted as a warrant as long as the student gave some indication of using it as explanation for the next claim. In some cases, this "warrant" did not actually support the claim, did not provide enough justification, or was a wrong statement entirely. These incorrect statements and implications also noted. About 30% of the *completely* warranted proofs had at least one incorrect implication. This statistic was even higher (at 43%) for the *mostly* warranted proofs. For those with *limited* warrants, it was 33%. Nevertheless, the majority of the warrants provided were correct.

The given scores also reflect the idea that warranted proofs generally score higher than those without warrants. Proofs with *complete* warrants on average had higher scores than other proofs. This is in agreement with Toulmin (1979), who claims that convincing arguments have warrants. Furthermore, results indicate that students provided more warrants in more familiar proofs (Integer proofs), as evidenced by the lower percentage of *mostly* warranted proofs in the Function category. This is in contrast to the conclusion at which Fukawa-Connely (2011) arrived, which was that students tend to provide more detail for topics they are unfamiliar with.

Backing

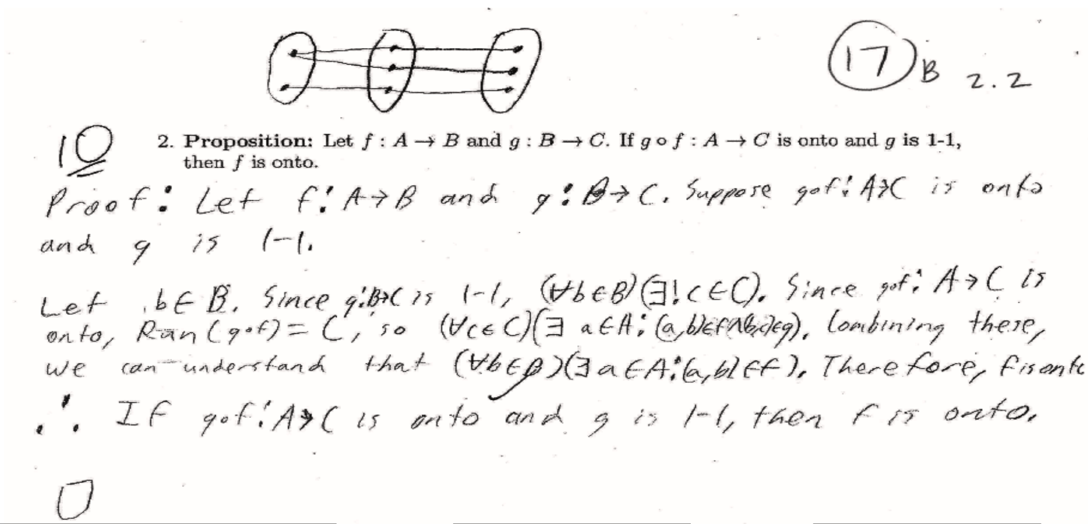
Overall, warrants found in Integer proofs contained more backing than those found in Function proofs. As with warrants, this seems to differ from Fukawa-Connely's (2011) research. With respect to backing, this data shows that students provided more detail in the familiar concept than in the unfamiliar concept.

In both the Function and Integer categories, more students had *limited* backing than *complete* backing. This may suggest that students were using self-evident backing as described by Simosi (2003) or that students struggled with identifying what information was needed to back a

…et $f: A \to B$ | $(\forall b \in B)$ | Then f is onto
…nd $g: B \to C$. | $\exists\, a \in A$
…uppose $g \circ f: A \to C$ | $s.t.\ (a,b) \in f$
…s onto and g is 1-1.
…et $b \in B$.

warrant. The fact that *completely* backed proofs had a slightly lower average score than the proofs with *limited* backing seems to support the idea that some pieces of backing are not crucial to a proof. While they do represent a more detailed proof, in many cases self-evident backing is appropriate. Rigorous and mathematically sound proofs usually contain some backing, but not always. In fact, there was one proof that received a score of 10 but provided no backing. This student's work is pictured below, followed by its schematic. The student chose to write many of his/her statements in set notation, which is valid, but tends to cut out the justifications.

Since $g: B \to C$ is 1-1 | Definition of onto
Since $g \circ f: A \to C$ is onto


17 B 2.2

10

2. Proposition: Let $f : A \to B$ and $g : B \to C$. If $g \circ f : A \to C$ is onto and g is 1-1, then f is onto.

Proof: Let $f: A \to B$ and $g: B \to C$. Suppose $g \circ f: A \to C$ is onto and g is 1-1.

Let $b \in B$. Since $g: B \to C$ is 1-1, $(\forall b \in B)(\exists! c \in C)$. Since $g \circ f: A \to C$ is onto, $Ran(g \circ f) = C$, so $(\forall c \in C)(\exists\, a \in A: (a,b) \in f \wedge (b,c) \in g)$. Combining these, we can understand that $(\forall b \in B)(\exists\, a \in A: (a,b) \in f)$. Therefore, f is onto.

$\therefore$ If $g \circ f: A \to C$ is onto and g is 1-1, then f is onto.

$\square$

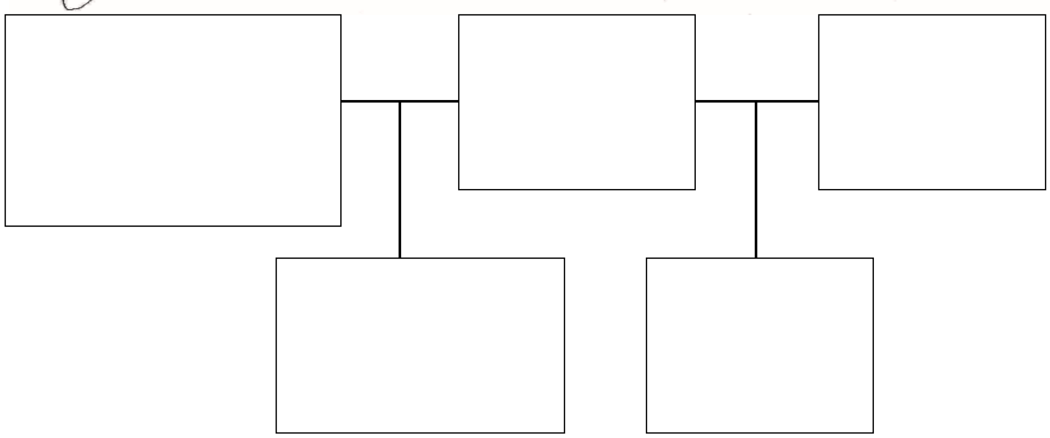


Figure 9: An Scn proof with a score of 10

Conclusion and Further Questions

There was a much wider variety of proof structures among student proofs in this research than anticipated. This suggests that students in the IBL class learned to take responsibility for their proof style and not solely rely on the authority of the instructor. Students were able to exercise creativity by presenting their own work to the class and critiquing work of their peers.

Since every student provided some level of warrant, students seemed to understand the importance of justification in mathematical proof. Backing, on the other hand, was less common. It was found that lack of backing did not significantly affect score, which may mean that implicit backing is acceptable in mathematical proof and not simply in oral argumentation. Would applying our coding scheme to problems from lecture-based classes or different IBL classes yield similar results?

After developing our scheme, we applied it to 16 proofs to see if any adjustments needed to be made before finalization. When coding the remaining proofs, we found that new situations arose and the coding scheme evolved. We took good note of every evolution, and had to go back to previously coded work to apply the most recent coding scheme. Would the scheme need to be further adjusted based on different problem types or different class levels?

Though Toulmin's (1979) original work was applied to non-mathematical argumentation, it has proved itself useful in analyzing the structure of student proof in a mathematical context. It is important for educators to value proof structure, for it is a key indicator of student understanding of mathematical proof.

References

Aberdein, A. (2006). The Uses of Argument in Mathematics. *Argumentation*, 287-301.

Arzarello, F., & Sabena, C. (2011). Semiotic and theoretic control in argumentation and proof activities. *Educational Studies in Mathematics*, 189-206.

Barrier, T., Mathe, A. C., & Durand-Guerrier, V. (2009). Argumentation and Proof: A Dsicussion About Toulmin's and Duval's Models. *CERME 6* (pp. 191-200). Lyon France: INRP.

Fukawa-Connelly (2011). Using Toulmin Analysis to Link an Instructor's Proof-Presentation and Student's Subsequent Proof-Writing Practices. 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 68-83. Portland, Oregon

Inglis, M., Mejia-Ramos, J., & Simpson, A. (2007). *Modelling Mathematical Argumentation: The Importance of Qualification.* New York: Springer.

Krummheuer, G. (2007). *Argumentation and Participation in the Primary Mathematics Classroom.* Orlando, Florida: Elsevier.

Rasmussen, C., & Stephan, M. (2008). A methodology for documenting collective activity. In A. E. Kelly, R. A. Lesh, & J. Y. Baek (Eds.). Handbook of innovative design research in science, technology, engineering, mathematics (STEM) education (pp. 195 - 215). Taylor and Francis.

Simosi, M. (2003). *Using Toulmin's Framework for the Analysis of Everday Argumentation: Some Methodological Considerations.* Netherlands: Kluwer Academic Publishers.

Toulmin, S., Rieke, R. D., & Janik, A. (1979). *An Introduction to Reasoning.* New York: Macmillian.

Correlations of students' ways of thinking about derivative to their success in solving applied problems

Shahram Shawn Firouzian
University of Maine
Shahram_Firouzian@umit.maine.edu

Students' understanding of derivative and their difficulties in solving applied problems have been the subject of rich research work. However little research has examined students' ways of thinking about derivative through the lens of their work on applied questions. The focus of this research is on whether relationships exist between students' ways of thinking about derivative and success on applied derivative problems. Survey data were used to look at students' multiple ways of thinking and their work on applied derivative problems. "Multiple ways of thinking" refers to two or more ways of thinking about derivative (e.g., slope of the tangent line at a point on a function or instantaneous rate of change). Findings indicate that students who demonstrate two or more ways of thinking about derivative were able to complete more steps of the applied problems.

Keywords: Applied problems, Derivative, Ways of thinking

Introduction and motivation

Among topics at the undergraduate level, derivative is regarded as an important foundational concept for further study in mathematics as well as in related areas and has been the object of a significant body of research (e.g., Ferrini-Mundy & Graham, 1994; Monk, 1988, 1994; Monk & Nemirovsky, 1994; White & Mitchelmore, 1996; Zandieh, 2000). In addition to these investigations of students' understanding and learning of derivative, researchers have also examined the role that student understanding of concepts connected to derivative play in understanding and learning of derivative. These "supporting concepts" (e.g., function, slope, limit) appear to interact with student understanding of derivative in important ways (Carlson, 1998; Monk, 1994; Monk & Nemirovsky, 1994; Robert & Speer, 2001). However, little is known about how students' conceptual understanding interacts with their work on applied problems such as those seen in mathematics, science and engineering courses.

For example, physics tasks about velocity, acceleration and radioactive decay require robust understanding of derivative. In chemistry, students learn how to calculate reaction rates, and in economics students work on maximizing profits, using marginal costs and marginal revenues, all of which are connected to rates of change. Many students have difficulties using knowledge and skills from one subject in other subjects. Physics instructors frequently note that students cannot use what they have learned in their mathematics classes in the context of physics problems (Basson, 2002; Bucy, Thompson, & Mountcastle, 2007).

In addition, even some of our best students do not completely understand important calculus concepts and when faced with an unfamiliar problem, have difficulty solving it (Carlson, 1998; Selden, Selden, Hauk, & Mason, 1999; Selden, Selden, Hauk, & Mason, 2000; Bezuidenhout, 1998).

The focus of this work is on the question of whether there is a relationship between students' multiple ways of thinking about derivative and their success in solving applied problems. "Multiple ways of thinking" refers to two or more ways of thinking about derivative (e.g., slope of the tangent line at a point on a function, instantaneous rate of change, etc.). Two different surveys were used to address these two domains of students' thinking. There were two studies conducted. In the first study, surveys were piloted separately in two different classes. After data were analyzed and preliminary findings were considered, both surveys were modified. In the second study, both surveys were deployed to the same subjects at the same time. Findings suggest that there is a correlation between students' having multiple ways of thinking about derivative and their completion of more steps of applied problems.

Theoretical Framework

Due to the complexity of the concept of derivative, we need to look at students' understanding of it from different perspectives. I used cognitive variability and strategy choice as described by Stiegler (2003) as features of students' thinking. Stiegler described how students at different ages use "multiple thinking strategies when solving problems of the same type" (p. 293). I describe these thinking strategies as ways of thinking about the concept of derivative following Harel's (1998) description of two categories of knowledge: ways of understanding and ways of thinking.

Harel (1998) suggests that students' ways of thinking are influenced by how they come to understand mathematical content. Therefore when we talk about students' conceptual understanding of derivative, I use "ways of thinking" to represent their concept images. Also to avoid defining knowledge in a static form and possibly neglecting the dynamic nature of knowledge and thinking, I use Harel's (1998) ways of thinking and understanding to define students' concept images of derivative.

Mathematics and science standards emphasize mastery of facts and procedures along with understanding of concepts (Hiebert & Lefevre, 1986). Procedural skills and conceptual understanding are usually highly correlated. However, conceptually oriented instruction has produced substantial gains in both kinds of knowledge (Siegler, 2003). Rittle-Johnson, Siegler, and Alibali (2001) define conceptual knowledge as "implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain" (p. 346-7). For topics such as derivative, using this definition means conceptual understanding of a multi-faceted idea. This, in turn, means that understanding of the concept of the derivative is displayed through the quantity of connections and relations between procedures, facts, representations and applications.

Zandieh (2000) used the notion of "concept image" to explore how students' ideas about derivative include a number of different representations of it. "Concept image" is defined by Tall and Vinner (1981) as "the total cognitive structure associated with the concept" (p. 152) which includes all the mental pictures and associated properties and images or multiple representations of derivative. Zandieh (2000) framed students' understanding of derivative in the multiple representations of graphical, verbal, physical, symbolical, and other. Abboud and Habre (2006) used graphical, numerical and symbolical views of derivative in assessing students' understanding of derivative. Kendal and Stacey (2003) used three representations of differentiation (graphical, symbolical, and physical) in creating what they called "differentiation competency framework."

In the application of this research to the analysis conducted in the present study, I combined categories developed by other researchers in some cases and added subcategories in others.

Students' solving applied derivative problems

In solving "real-world" problems, Tall (1991) wrote that the given problem is first translated from the context to the abstract level of calculus, the abstract problem is then solved, and the solution is translated back to the context. The first step obviously calls on students' conceptual knowledge of variables, algebra skills, and calculus concepts because it depends on the identification not only of the appropriate concepts in the given context but also of the relationships among them. For example, the identification of appropriate concepts might involve the selection of one or more symbolized variables from among several concepts.

White and Mitchelmore (1996) divided 40 first year college students into four groups and conducted four versions of four different applied derivative problems (related rate and optimization problems) from high level of translation requirements to the lowest level. They showed that initially students could not apply their knowledge of calculus, even though they had seen similar items before. After several extra sessions, average performance still only exceeded 50% and later, almost all could correctly symbolize, but only slightly more than 60% could then proceed to a correct solution.

The results from White and Mitchelmore's (1996) maximization question showed that most students either could not identify or symbolize an appropriate variable by translating one or more quantities in the item to an appropriate symbolic form or if they could, were unable to make use of their definitions.

Roorda, Vos, and Goedhart (2007) developed a framework that focused on representations as part of understanding derivatives, but included applications as well. They did not, however, investigate whether there is any relation between students' having multiple ways of thinking about derivative and how well they do on applied problems.

In applying their knowledge of function and derivative to graphical questions, Stahley (2011) examined students' qualitative understanding of the relationship between the graph of a function and the graphs of its first and second derivatives. The goal was to understand the thought processes of students as they solved or attempted to solve questions about derivatives. He also reported on the common and uncommon inaccurate ideas students displayed as they completed conceptually based tasks about graphs of functions and their derivatives. Berry and Nyman (2003) showed that many students can differentiate an algebraically presented function, but differentiating or sketching the derivative of a graphically presented function requires dissimilar thought processes and analytical skills. According to Monk (1988), many students have point-wise understanding of function with less across-time understanding where "across-time" means being aware of the behavior of the original graph and understanding the relationship between it and its derivative sketch. There has not been much work done on the correlations of students' multiple ways of thinking and how well they do on graphical derivative questions.

Materials

Study I

In the first study, two surveys were analyzed separately. One survey with five questions was about students' ways of thinking about derivative and is referred to as the *derivative survey*. The second survey contained two applied derivative problems and is referred to as the *applied question survey*. As explained later, this first study focused on expanding our understanding of the students' concept images of derivative, and their difficulties on applied problems separately.

Some questions on the derivative survey were based on existing research and others were created by the researcher. Below are three sample questions from the survey.

1. You are talking with someone who just started high school. In a sentence or two explain to them what is meant by the derivative of a function. (Feel free to use any graph, symbols, or words in your explanation.)

Question one was designed similarly to the work of Zandieh (2000). As Zandieh explained, for a concept as multifaceted as derivative it is not appropriate to ask whether or not a student understands the concept. One should ask instead for descriptions of students' understanding of derivative to determine what aspects of the concept students know and the relationships students see between those aspects.

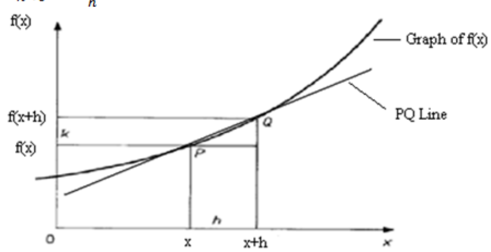
2. If y is a function of x, explain in words the meaning of the equation $\frac{dy}{dx} = 5$ when x = 10.

This question was borrowed from Kendal and Stacey (2003). Using their differentiation competency framework they traced students' learning of three representations. This question was used to probe students' understanding of derivative in terms of numerical values. In their findings, this question had one of the lowest numbers of correct responds.

3.

Here is a definition of the derivative:

$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$

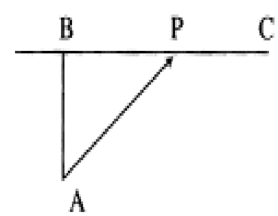


Why does the formula give us the derivative of the function? Provide an explanation for your answer.

This question was designed by the researcher in collaboration of experts in the field to probe students' symbolical understanding of the derivative when it is given as a limit of a difference quotient. By including the graph we were hoping the students' graphical understanding of derivative would also be prompted.

The second survey (applied question survey) was aiming at students' thinking about applied problems. Both questions were borrowed from White and Mitchelmore (1996). One question was about rate of change and one was about maximization. These questions were selected because they required the most translation, meaning that the questions require the definition of new variables and symbolic expression of relations between variables, and selection of a calculus concept and its expression in symbolic form. However, one feature of these questions is a low number of variables and that makes symbolizing relatively straightforward.

1. If the edge of a contracting cube is decreasing at a rate of 2 centimeters per minute, at what rate is the volume contracting when the volume of the cube is 64 cubic centimeters? (Provide an explanation for your answer.)
2. The diagram shows a straight road BC running due east. A vehicle is in open country at A, 3km due south of B. It must reach C, 9km due east of B, as quickly as possible. The driver can travel at 80kph in open country and 100kph on the road. Assuming the car proceeds through open country to some point P, and then along the road, what is the distance of P from B so that the journey to C takes the shortest time possible? (Provide an explanation for your answer.)



Study II

For the second study, the correlation between students' multiple ways of thinking about derivative and their success on applied problems was investigated. Both surveys from study I were modified before being used in the second study. The derivative survey was modified to include a set of multiple-choice questions to help ensure that students had opportunities to describe all their ways of thinking about derivative. Students were asked whether the statement described the idea of a derivative and they had three answer choices: yes, no or not sure. Questions included, for example, were "How much one quantity change with respect to another quantity, "Slope of any point of the graph," and "How much faster something is going." These multiple choice questions were designed based on other researchers' findings and analysis of students' work in the first study. The wording used was as close as possible to students' written from the first study. To help ensure that we captured all the possible ways students might have thought about derivative, one last question was also included: "If any other ideas come to mind besides what you marked above, please write them here."

The modified derivative survey I also included a graphical question, asking students to draw first and second derivatives of a function (Fig. 1). The question (below) was borrowed from Stahley (2011) the purpose of using it was to look at possible connections between students having multiple ways of thinking about derivative and success on graphical derivative questions.

Directions: Sketch graphs of the first and second derivatives of the given function.

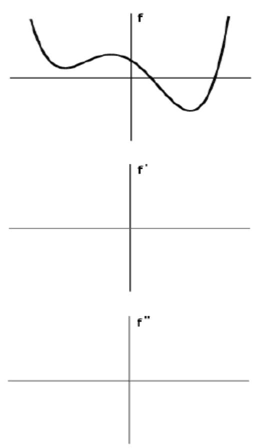
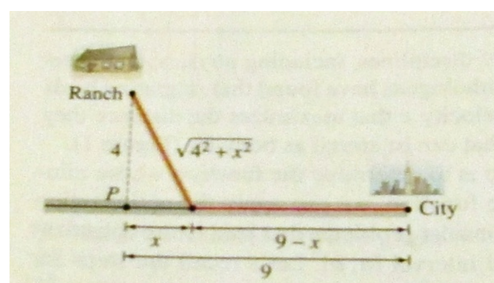


Figure 1. Graphical applied question

As discussed in the analysis of the results, for the second study, a modified version of the applied question survey was used. The first question was the same and second question was modified to include definitions of the variables.

2. Cowboy Clint wants to build a dirt road from his ranch to the highway so that he can drive to the city in the shortest amount of time (Figure below). The perpendicular distance from the ranch to the highway is 4 miles, and the city is located 9 miles down the highway. Where should Clint join the dirt road to the highway if the speed limit is 20 mph on the dirt road and 55 mph on the highway?



Data Collection Methods

The participants in these studies were students in differential calculus and multivariable calculus classes at a public university in the northeastern United States.

Study I

For the first study, both surveys were administered separately in two different classes. The derivative survey was given to 125 students in a differential calculus class during the last two weeks of the semester, thus it was after they had been introduced to the derivative. The university offers only one type of calculus class and as a result, students had a range of backgrounds and were in various majors (including physical sciences, mathematics, engineering, biology, education, and social sciences).

The applied question survey was given to 51 students in a multivariable calculus class. These students had completed the differential calculus course with at least a grade C. Students were explicitly asked to describe their solution or to explain why they could not solve it.

Study II

Once both surveys of study I were modified, they were given to 52 students in a differential calculus class. Students completed and turned in the derivative survey and then they were given the applied question survey. For the applied question survey, the formulas for the speed and Pythagorean Theorem were provided. Students were again asked to explain their solutions or to say why they were unable to complete the problem.

Data Analysis Methods

All coding and analysis were done on students' written work. Some coding was based on other researchers' methods (Abboud & Habre, 2006; Kendal & Stacey, 2003; Zandieh, 2000) and some coding was developed using grounded theory (Glaser and Strausss, 1967).

A set of categories for students' ways of thinking were defined by the researcher. This set was based on similar categorization schemes used by other researchers (e.g., Kendal & Stacey, 2003; Zandieh, 2000):

- **Symbolical**: the formal definition of derivative limit of difference quotient.
- **Graphical**: uses slope of the tangent line at a point on the graph or function.
- **Numerical** This category refers to when the students use numerical values to describe what a derivative is.
- **Verbal**: when students use "instantaneous rate of change."
- **Physical:** when students use velocity, speed or acceleration to describe the derivative.
- **Procedural:** when students talk about power rule or actually writes an example of taking the derivative as a way of explaining it, for instance: $f(x)=x^2$ so $f'(x)=2x$.
- **Other:** when students use area under the graph or accumulation to explain the derivative.

As seen on the figure 2, when the students think of derivative in ways similar to the graphical numbers 1 and 2, they are considered (complete) Graphical. However if their definitions do not include all the things needed to be a complete description of graphical way of thinking they are considered "Graphical (partial)" (categories 3 to 7 in the Fig. 2).

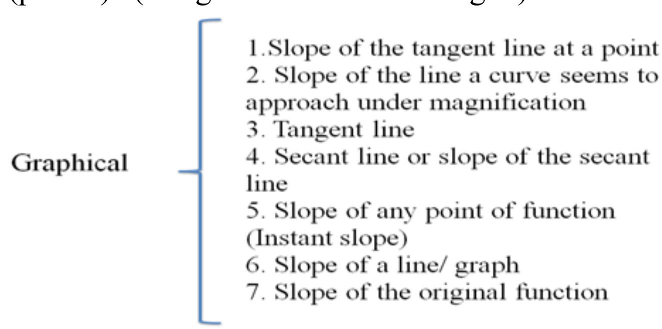


Figure 2. Categories for graphical ways of thinking about derivative

We find the same scenario for the Verbal way of thinking category. When students used "instantaneous rate of change of a function" to define the derivative, it was categorized as a (complete) Verbal way of thinking about derivative. As seen on the below image, students also had other ways of thinking that were Verbal but did not fully capture all the ideas. These are referred to as "Verbal (partial)" categories. For instance, when they talk about "how much something is going faster" their way of thinking about derivative is also categorized verbal, except that it is "Verbal (partial)" (Categories 2 to 7 in Fig. 3).

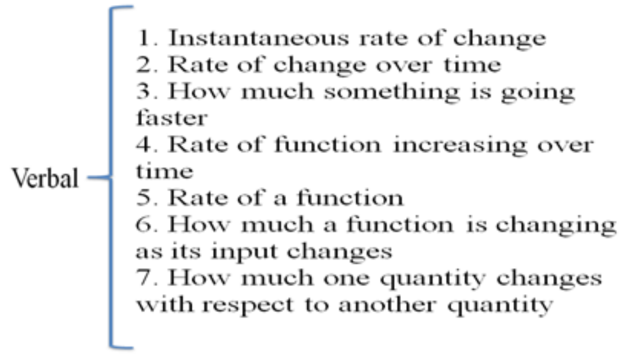


Figure 3. Categories for verbal ways of thinking about derivative

For the applied problems, analysis methods from White and Mitchelmore (1996) were used. There were three main categories of error, in all of which the variables are treated as symbols to be manipulated rather than as quantities to be related. According to White and Mitchelmore, the dominant error centered on students' inabilities to correctly use the specific value for volume (i.e., V=64) in conjunction with the equation for volume (i.e., $V = x^3$). They also noticed another common difficulty, called "manipulation focus," in which students' base decisions about which procedure to apply on the given symbols and ignore the meanings behind the symbols.

For study II, the researcher created a rubric of nine steps (below) based on Tall's (1991) interpretation of steps taken by students to solve the problem 1 of the applied question survey (Fig. 4). By dividing the process of solving the applied problems to these nine steps, the researcher also hoped to track all possible errors they made. This rubric is similar to Martin's (2000) classification of steps in solving geometric related-rates problems except our questions are context based.

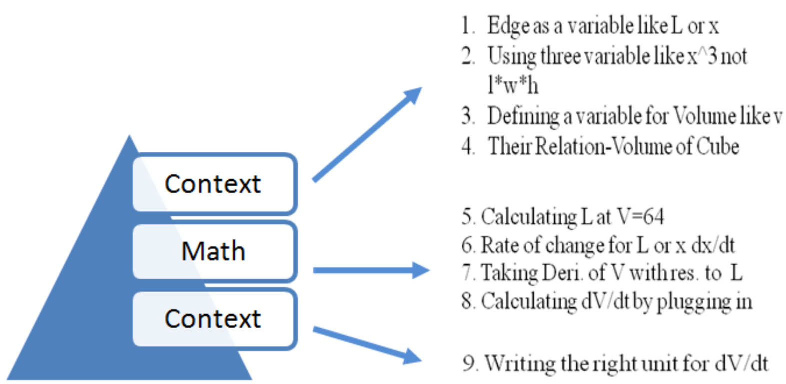


Figure 4. Nine steps rubric for question 1 of the applied

The second question of the applied question survey was inspired by White and Mitchelmore's (1996) maximization question. Another common error they identified beside manipulation focus was what they call "x,y syndrome. " This is when the procedure for maximizing a function has apparently been remembered in terms of the symbols used when it was first learned, rather than as a process based on the rate of change of the function.

When it comes to students' difficulties with graphical questions where they are asked to sketch derivative graphs, Stahley (2011) demonstrated that students have three main difficulties when attempting to sketch the first and second derivatives of a given function. First, students have the tendency to sketch derivatives that represent opposite (meaning reflection over the horizontal axis) behavior of the original function. Second, proceeding left to right, many first derivative sketches begin correctly, but somewhere along the sketching process students become confused and the sketch ends up being incorrect. Lastly, many students express confusion about the second derivative and often sketch a linear function for the second derivative independent of the degree of the original function.

For the Study II, using Stahley's definition of correct answers, and using the macimum, minimum, and crticial points, I created a six step rubric for the first derivative and four steps rubric for the second derivative to quantify the students' work on the graphical question. Once the students had the first derivative graph sketched, they then need to use that in order to sketch the second derivative. In order to complete the first derivative the students need to look at the critical points, max and mins. For each point, students need to look at the graph before the point, at the point and after the point which adds up to six steps from the original function to the first derivative. To sketch the second derivative the students need to look at the max and mins and the graph behaviors before and after them which add up to four steps.

Findings: Study I

Students' ways of thinking about derivative

Findings confirmed those from existing research where the majority of students give verbal or graphical definitions of derivative. However within those categories of thinking we find the majority think of derivative in graphical (partial) or verbal (partial) ways.

110 of 125 students used graphical ways of thinking to define derivative, however only 19 of 110 used the full graphical definition (Categories 1 and 2 of Fig. 2) and 95 of 110 (or 76%) defined the derivative using graphical (partial) (Categories 3 to 7 of Fig. 2).

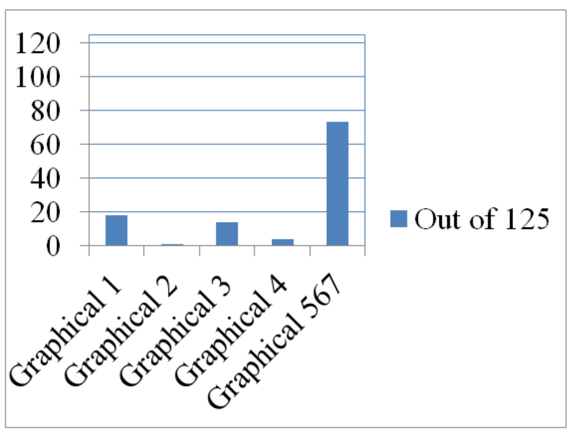


Figure 5. Number of responses in the categories of graphical ways of thinking about derivative based on the Fig. 2

More than 65% of the students defined derivative using a verbal way of thinking (107 out of 125). Only 18 of the 107 defined the derivative using the full verbal (category 1 on Fig. 3) definition as the instantaneous rate of change of a function.

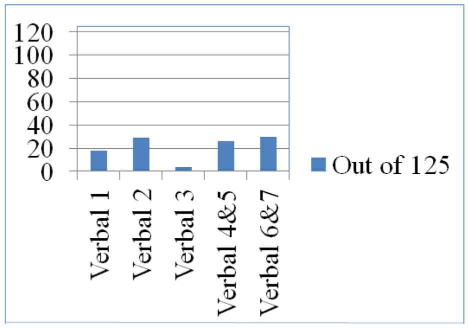


Figure 6. Number of responses in the categories of verbal ways of thinking about derivative based on the categories of Fig. 3

There were other ways students defined derivative (Fig. 7). Only 3 out of 125 students defined the derivative using symbolical way of thinking even though they were given the definition in the third question. Out of 125 students, 11 used physical ways of thinking about derivative.

After graphical and verbal ways of thinking about derivative, procedural was the next most common way of thinking about derivative. More than 24% of the students described a procedure (31 out of 125) as a way to define derivative. A very small number of students used numerical values in defining the derivative even though the second question was aiming at their understanding of derivative using numerical values. Out of 125 students only 9 defined the derivative using numerical values.

Many students also used accumulation over time to define the derivative. These, as well as blank answers, were assigned to the other category. Only 9 students out of 125 used these ways of thinking to define the derivative.

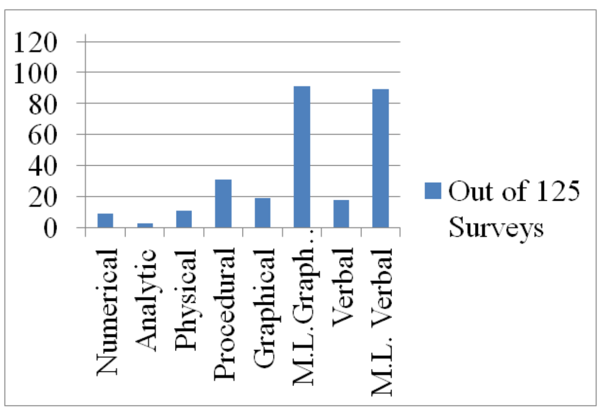


Figure 7. Accumulated results of students' ways of thinking about derivative

Findings: Students' success on the applied problems

There already exists research on students' difficulties with defining variables (Schoenfeld & Arcavi, 1988; McNeil, Weinberg, Hattikudur, Stephens, Asquith, Knuth & Alibali, 2010). For the first question, borrowed from the related rate problem of White and Mitchelmore, they pointed out that in the last two collections of their four collections, almost all the students could correctly symbolize, but only slightly more than 60% of these students could then proceed to a correct solution. They found that dominant error centered on students' inability to correctly use V=64. Sample responses included: some students left the answer at $-6x^2$, indicating that the 64 had not been used. Some students gave two answers, one for $V=x^3$, and one for V=64. They concluded that both errors had the same source: the inability to reconcile the variable expression $V=x^3$ with particular value V=64.

Out of the 51 surveys, only 10 people got the correct answer for this question. 41 out of 51 students could not answer it correctly. These findings are consistent with those from other researchers. The following bar graph summarizes the correct results for each step of the rubric.

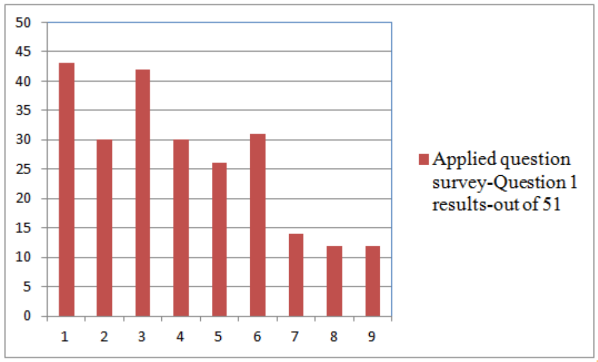


Figure 8. Bargraph of applied problem survey question 1 results based on the rubric of Fig. 4

On the applied problem survey, no student answered the second question correctly. After analyzing students' written work, four main categories of difficulties were determined:

1. Wrote "Could not remember" how to set it up or no answer (15/51)
2. Attempted but no defining any variables for any quantities or wrong definitions of variables (25/51)
3. Correct Variables but wrong modeling (8/51)
4. Correct Translation but wrong calculus- x-y syndrome-Manipulation focus (3/51)

A majority of students could not translate from the question context to the calculus. On the applied problems we wanted to focus on the calculus part of the question but students could not go beyond the algebra steps.

There were few issues with the results of study I. We were not sure that the written responses displayed all that the students knew. For the first survey, using students' written work from study I, the researcher prepared a set of multiple-choice questions that addressed all the ways of thinking about derivative students had offered during study I. This helped ensure that students were given every possible option to reveal all their ways of thinking about derivative.

For the applied problem survey, in order to study the students' ways of thinking about the derivative and its translation in the applied problem, and since majority of their problems were in their unsuccessful definition of variables, the second question was modified to include variables.

Study II

Modified surveys were rerun simultaneously. Using the same coding from study I, we found the results in image below. 78% of the students who described the derivative with two or more ways completed more steps of the first applied problem (related rate problem).

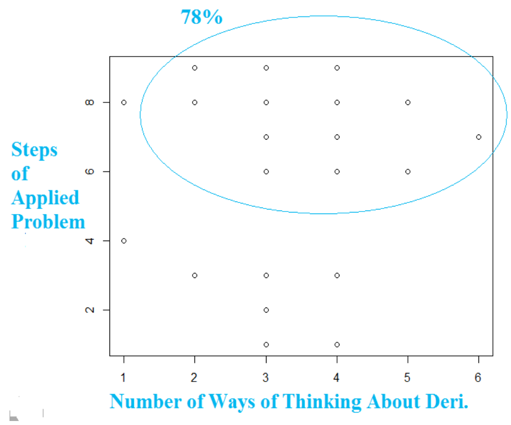


Figure 9. Image-Relation of students' having multiple ways of thinking about derivative and their success in the applied derivative problem

On the second question of the second survey, nobody got the correct answer. 88% of the students could not define the relation between the variables even thought the speed and Pythagorean formulas were given to them.

On the graphical application of derivative problem, Stahley (2011) found that 70% of the students sketched the first derivative correctly and 64% got the second derivative correctly. These results are similar to our findings: 74% of the students completed the first derivative correctly and 73% of those 74% (54% of the total students) sketched the second derivative correctly.

Interestingly, out of the 74% students who sketched the first derivative correctly, 92% of them had more than two ways of thinking about derivative. And the students who sketched the second derivative correctly all had multiple ways of thinking about derivative.

Conclusions & Implications

By expanding existing research results on students' ways of thinking about derivative and their difficulties with the applied derivative problems, we have unraveled a hidden correlation between students having multiple ways of thinking about the derivative and how well they complete the applied problems.

Our findings confirm some of the full category of students' ways of thinking about derivative as defined by other researchers (Zandieh, 2000; Abboud & Habre, 2006). We also confirmed that the majority of students had rate of change as their most prominent interpretation of what a derivative was. Slope of the tangent line as the second most common interpretation, and just one out of 34 students stated the formal definition of derivative. However, very few explicitly discussed the procedural way of thinking about derivative which is when the students treat the derivative as procedure. And as we saw, students might not think of derivative fully as the slope of the tangent line at a point on the function or graph and they might think of it as "slope of the tangent line" without linking it to function. This was our categories of verbal (partial) and graphical (partial) ways of thinking about derivative.

On the second question of the first survey, Kendal and Stacey (2003) showed that a small percentage of the students could interpret the questions and majority could not even define variables for the different quantities. Our findings also confirmed their results however their second, third, and forth versions of their original optimization question was different from our modified version of the applied problem survey. They removed the context of the question and made it purely mathematical and students could solve them purely with procedural understanding of derivative. Our second version of the question preserved the original context however defined the variables for the students even though none of the students could solve the problem correctly. Giving contextual applied problem from other discipline, enhance our students' translation skills in defining the variables and their relationships.

Graphical questions can perhaps be used for the students to transfer their multiple representational knowledge of derivative. The nature of having multiple ways of thinking about derivative and students better success on graphical questions shows us the importance of graphical applied problem as a great instructing tool.

Study II was only done with 34 students and to support our hypothesis we need to run the surveys with a bigger population. The nature of this connection needs to be explored deeper with clinical interviews. Further analysis on the data can also reveal further interconnectedness between different ways of thinking and how successful they are on different applied problems.

Calculus instructors should provide the students with opportunities to explore different ways of thinking about derivative in order for them to utilize that learning in different contexts and on different applied problems. Perhaps reviewing some of the basic algebra skills in order to better understand the concept of derivative and its applications will prove to be useful for the students in the long run. Just teaching the students rote memorization of statements that describe the derivative or just procedural methods appears not to help them advance their knowledge and skills to solve unfamiliar problems.

In the same way I created a multiple choice questions to assess all their possible ways of thinking about derivative, one could use the same survey to see their partial ways of thinking about derivative. That information could inform instruction so it gives them opportunities to complete their understanding. A lot of ways to describe derivative historically came from other discipline contexts like physics, chemistry, etc. Perhaps introducing the derivative from the applied context will help our students develop the different ways of thinking about derivative in the context instead of having surface understanding of it.

The importance of thinking about derivative in multiple ways which ultimately enhance their success in applied problems requires us to provide our teachers with more opportunities to learn about students' difficulties and thinking about derivative and its applications as well. The partial understanding and different ways of thinking about derivative will help instructors know how their students might think about derivative. The common errors and difficulties the students had with applied problem can help the instructors to focus on particular students' difficulties. The next phase of this project will be exploring teachers' integrated mathematical and scientific knowledge for teaching. By having students' ways of thinking about derivative and applied problems, the researcher will be exploring this untouched domain of teachers' knowledge.

Acknowledgement

I would like to extend my sincere appreciations to my mentor and friend, Professor Natasha Speer, my wife Parisa, and all of my colleagues in University of Maine-Mathematics Education Research Group.

References:

Abboud, M., & Habre, S. (2006). Students' conceptual understanding of a function and its derivative in an experimental calculus course. *Journal of Mathematical Behavior,* Vol. 25, No. 1, (pp. 57-72)

Bezuidenhout, J. (1998). First-year university students' understanding of rate of change. *International Journal of Mathematical Education in Science and Technology, 29(3),* 389-399.

Basson, I. (2002). Physics and mathematics as interrelated fields of thought development using acceleration as an example. *International Journal of Mathematics Education in Science and Technology*, *33*(5), 679−690.

Berry, J. N., M. (2003). Promoting students' graphical understanding of the calculus. *Journal of Mathematical Behavior, 22*, 481-497.

Bucy, B. R., Thompson, J. R., & Mountcastle, D. B. (2007). Student (mis)application of partial differentiation to material properties. *Proceedings of the 2006 Physics education Research Conference, American Institute of Physics,* 157-160.

Carlson, M. P. (1998). A cross-sectional investigation of the development of the function concept. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education. III. CBMS Issues in Mathematics Education* (pp. 114-162). Providence, RI: American Mathematical Society.

Ferrini-Mundi, J. & Graham, K. (1994). Research in calculus learning: understanding of limits, derivatives and integrals. In J. J. Kaput & E. Dubinsky (eds.), *Research Issues in Undergraduate Mathematics Learning*, MAA Notes 33, (pp.31-45). Washington DC: MAA.

Glaser, B. & Strauss, A. (1967), *The Discovery of Grounded Theory.* Chicago: Aldine Publishing Company.

Harel, G. (1998). Two dual assertions: The first one learning and the second teaching (or vice versa). *American Mathematical Monthly, 105(6), 497-507.*

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–28). Hillsdale, NJ: Erlbaum.

Kendal, M., & Stacey, K. (2003). Tracing Learning of Three Representations with the Differentiation Competency Framework. *Mathematics Education Research Journal*, *15*(1), 22-41.

Martin, T. (2000). Calculus Students' Ability to Solve Geometric Related-Rates Problems. *Mathematics Education Research Journal, 12(2), 74-91.*

McNeil, N. M., Weinberg, A., Hattikufur, S. Stephens, A. C., Asquith, P. Knuth, E. J., & Alibali, M. W. (2010). A is for "Apple": Mnmonic Symbols Hinder the Interpretation of Algebraic Expressions. *Journal of Educational Psychology, 102(3),* 625-634

Monk, G. S. (1988). Students' Understanding of Functions in Calculus Courses. *Humanistic Mathematics Network Newsletter, #2, March 1988*

Monk, G. S. (1994). Students' Understanding of Functions in Calculus Courses. *Humanistic Mathematics Network Journal, 9*, 21-27.

Monk, S., & Nemirovsky, R. (1994). The case of Dan: Student construction of a functional situation through visual attributes research in Collegiate Mathematics Education,4 , 139-168.

Orton,A. (1984). Understanding rate of change. *Mathematics in School*, 13(5), 23-26.

Rittle-Johnson, B., Siegler, R., & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346–362.

Robert, A., & Speer, N. (2001). Research into the teaching and learning of calculus/elementary analysis. In D. Holton (Ed.), *The teaching and learning of mathematics at university level: An ICMI Study* (Vol. 7, pp. 255-273). Dordrecht: Kluwer Academic.

Roorda, G., Vos, P. & Goedhart, M. (2007). Derivatives in applications: How to describe students understanding. In G. Kaiser, F. Garcia & B. Sriraman (Eds.), *Proceedings of the WG on 'Mathematical Modeling and Applications' 5th Conference on European Research in Mathematics Education (CERME-5).* Nicosia, Cyprus: University of Cyprus.

Selden, J., Selden, A., Hauk, S. & Mason, A. (1999). *Do Calculus Students Eventually Learn to Solve Non-Routine Problems*. Cookeville, TN: Tennessee Technological University.

Selden, A., Selden, J., Hauk, S., & Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? In A. H. Schoenfeld., J. Kaput, & E. Dubinsky, (Eds.), *Issues in mathematics education: Vol. 8. Research in collegiate mathematics education.* IV (pp. 128-153). Providence, RI: American Mathematical Society.

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.

Siegler, R. S. (2003). Implications of cognitive science research for mathematics education. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 289–303). Reston, VA: National Council of Teachers of Mathematics.

Stahley, J. R. (2011). *Students' qualitative understanding of the relationship between the graph of a function and the graphs of its derivative.* Retrieved from Electronic Theses and Dissertation. Paper 1609.

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

Tall, D.O. (1985). Understanding the calculus. *Mathematics Teaching,* 110, 49-53.

Tall, D.O. (1991). Intuition and rigour: the role of visualization in the calculus. *Visualization in Teaching and Learning Mathematics* (ed. S. Cunningham & W. S. Zimmermann, MAA Notes No. 19, (pp.105-119). Washington DC: MAA.

White, P., & Mitchelmore, M. (1996). Conceptual knowledge in introductory calculus, *Journal for Research in Mathematics Education, 27,* 79-95.

Zandieh, M. J. (2000).A theoretical framework for analyzing student understanding of the concept of derivative. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *CBMS Issues in Mathematics: Research in Collegiate Mathematics Education, IV*(8), 103–127.

A CASE STUDY ON A DIVERSE COLLEGE ALGEBRA CLASSROOM: ANALYZING PEDAGOGICAL STRATEGIES TO ENHANCE STUDENTS' MATHEMATICS SELF-EFFICACY

Michael Furuto and Derron Coles
Oregon State University

Shifting demographics show America rapidly diversifying, yet research indicates that an alarming number of diverse students continue to struggle to meet learning outcomes of collegiate mathematics curriculum. Consequently, recruitment and retention of diverse students in STEM majors is a pervasive issue. Using a sociocultural perspective, this study examined the effect of two pedagogies (traditional instruction and cooperative learning) in a diverse College Algebra course on enhancing students' mathematics self-efficacy. Particular attention was paid to investigating the role interaction and discourse play in facilitating learning, improving conceptual understanding, and empowering students to engage in future self-initiated communal learning. The goal was to develop an effective classroom model that cultivates advancement in knowledge and enculturation into the STEM community, culminating in a higher retention rate of diverse students in STEM disciplines. Results indicate that a hybrid model encompassing both traditional instruction and cooperative learning successfully enhances students' self-efficacy.

Key words: Diversity, Pedagogy, College Algebra, Sociocultural Theory, Self-Efficacy

Introduction

Research indicates that an alarming number of diverse[1] students struggle to meet learning outcomes in mathematics (National Center for Education Statistics, 2009; Trends in International Mathematics and Science Study [TIMSS], 2007). This is a critical issue because according to the U.S. Census, the racial and ethnic minority population of the U.S. continues to expand rapidly while the percentage of this population who avoid or fail to complete degrees in Science, Technology, Engineering, or Mathematics (STEM) disciplines remains steady. Deficiency in mathematics competency is a major cause of difficulties in recruitment and retention in STEM (Huang, 2000). For the U.S. to remain competitive in the world market of global technology, an increase in minority STEM leaders is essential and must be addressed.

The achievement gap in mathematics for diverse students begins in K-12 with ripple effects experienced throughout secondary education (Bettinger & Long, 2006). Although mathematics educators (Boaler & Staples, 2008; Gutstein, 2003) have conducted extensive research on this issue at the K-12 level, significantly less has been completed at the college level. Understanding college student development is necessary for implementing effective scaffolding at all levels of education. We seek to contribute by examining mechanisms for increasing mathematics self-efficacy (Bandura, 1977), or one's confidence in completing a task, of diverse college students.

Literature Review

Research studies demonstrate that students with high self-efficacy score higher academically (Multon, Brown, & Lent, 1991). Hackett and Betz (1989) assert that even when variables such as

[1] The NCATE (2008) defines *diversity* as: "differences among groups of people and individuals based on ethnicity, race, socioeconomic status, gender, exceptionalities, language, religion, sexual orientation, and geographical area."

mathematics aptitude, gender, and anxiety are controlled, self-efficacy beliefs are predictive of students' choice of major and academic performance. We take the perspective that building high self-efficacy will benefit diverse students not only in the short-term (i.e. passing classes), but have significant long-term impacts as well. Diverse students with high self-efficacy are more likely to persist to graduation in STEM fields, which could ultimately lead to diversifying STEM disciplines. To advance diversity in STEM, teachers play an instrumental role in implementing effective pedagogical strategies and establishing a supportive learning environment.

Over the years, teachers have implemented various pedagogical methods in classrooms to heighten learning. Commonly employed pedagogies include variations on *traditional instruction* (where teachers exclusively lecture and students work independently). Through independent learning, students can increase their self-efficacy by solving problems without aid. In contrast, *cooperative learning*[2] or peer interaction in groups has since emerged as one of the prominent methods in developing learning and confidence. The NCTM (1991) asserted, "Whether working in small or large groups, [students] should be the audience for one another's comments – that is, they should speak to one another, aiming to convince or to question peers" (p. 45). Peer interaction provides students opportunities to work with classmates, explain solution methods, and construct knowledge while participating in a community of practice (Lave & Wenger, 1991).

The extent of student interaction and discourse is the most significant difference between the two learning contexts. The NCTM (1991) describes discourse as teachers facilitating and students engaging in "ways of representing, thinking, talking, agreeing, and disagreeing" (p. 34). According to the NCTM, communication is necessary in mathematics education. Since the NCTM published new material promoting discourse, many educators have implemented the usage of discourse into the curriculum. Research indicates that learning through discourse is central to helping students develop their mathematical understanding. Discourse has helped students synthesize mathematics (Chaplin, O'Connor, & Anderson, 2009), and improve their cognitive development (Cobb, Wood, & Yackel, 1993). We hypothesize that working in collaborative groups will promote student engagement in discussions, expand learning by combining each member's knowledge base, and encourage students to teach and learn from their classmates, which can ultimately increase self-efficacy such that students persist in STEM.

Research Questions

This study explores the following research questions:

1) What do individual students perceive to be the benefits and challenges of traditional and collaborative pedagogies, particularly with respect to enhancing their mathematics self-efficacy?
2) What evidence exists indicating the role of discourse in improving diverse students' abilities to learn and do mathematics?
3) What are the implications of addressing diverse student perspectives when implementing either pedagogy?

Theoretical Framework

We have framed this study using a sociocultural theoretical perspective. At the heart of cooperative learning and student discourse is the acquisition of knowledge within socially and culturally shaped contexts. Vygotsky (1978) described sociocultural theory as an ongoing

[2] *Cooperative learning*, also known as collaborative learning, will be defined as "small groups of learners working together as a team to solve a problem, complete a task, or accomplish a common goal" (Artz & Newman, 1990, p. 448).

process where learning is facilitated by interaction within social and cultural contexts. Vygotsky stressed that to advance learning, the human mind must operate in a social environment where "every function in the [student's] cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological) and then inside the child (intrapsychological)" (p. 57). In our research, we focus on Vygotsky's two levels of learning, but interpreted by the students experiencing the learning firsthand. We use the students' perspectives to investigate how students use the cooperative (highly interpsychological) and traditional (highly intrapsychological) learning contexts to navigate around their unique barriers to knowledge acquisition. In the process, we intend to gain a student-focused insight into the advantages and disadvantages of both pedagogies.

Proficient social interaction, whether peer-to-peer or teacher-to-student, through discourse provides students with abundant opportunities to cultivate learning. In Bauersfeld's (1980) study of improving mathematics comprehension, he declared, "Teaching and learning mathematics is realized through human interaction" (p. 35). We acknowledge that through the lens of sociocultural theory, both the traditional and cooperative pedagogies have social and environment-dependent components. The zone of proximal development or ZPD (Vygotsky, 1978) defined as the difference between advances in knowing brought about by independent knowledge acquisition versus socially directed learning should be detectable despite the pedagogy. Even the most passive knowledge transition lecture style could be considered a situation where students are developing through guidance and could result in a measurable increase in learning associated with the ZPD. However, the degree to which the ZPD is affected likely depends on the level to which either pedagogy is situated to be an effective venue for idea exchange. Our analysis looks specifically at the classroom culture as shaped by individual student identities and most significantly by pedagogy. We assert it is the pedagogy creating the culture of the classroom that sociocultural theory leads us to believe impacts learning opportunities. The degree to which the students interact with each other and the teacher under the "constraints" of a pedagogically driven classroom culture is an additional topic of interest.

Methodology

Given our goal of comprehending the highly complex phenomena of teaching and learning mathematics, we collected data using multiple sources in a mixed-method research design. We studied one College Algebra course with diverse students (n = 10) that took place during the summer term of 2012. Data were collected at a medium sized public university in the Pacific Northwest. As illustrated in Table 1.1 below, the students in this class were all diverse, including older-than-average, veteran, single parent, ethnic and racial minority, rurally isolated, 1st generation in college, and international students. The class was evenly split with five male and five female students. Three of the ten participants were Hispanic, two international (one from China and one from Saudi Arabia), one Asian American, and four White. Out of the four White students, two were older-than-average and veterans, and the other two were 1st generation college students. The students and their demographic information are summarized in Table 1.1.

Table 1.1 – *Background Information on Participants*

Name	Gender	Age	Ethnicity	Year in School	Major
Adam[3]	Male	26	White	Junior	Fisheries and Wildlife Biology

[3] To maintain conditions of anonymity, all names used throughout this study are pseudonyms.

Angela	Female	29	Hispanic	Junior	Mechanical Engineering
Cassie	Female	21	White	Senior	Merchandising Management
Chelsea	Female	22	Asian American	Junior	Merchandising Management
Dante	Male	22	Hispanic	Sophomore	Agricultural Business
Jason	Male	25	White	Senior	New Media Communications
Jin	Male	20	Chinese	Freshman	Biology
Juanita	Female	22	Hispanic	Senior	Apparel Design
Rahul	Male	27	Saudi Arabian	Junior	Electrical Engineering
Sara	Female	19	White	Freshman	Exercise and Sport Science

While this study is a mixed-method design, in order to obtain the most valid and reliable data to specifically address the research questions, qualitative research is the preferred method. Regarding case studies, qualitative research has proven to be an efficient method of discovering rich, in-depth understandings of teaching and learning mathematics. To support the qualitative data and strengthen findings, we administered Hackett and Betz' (1989) Mathematics Self-Efficacy Scale (MSES) as a pre/post survey, which has a reliability coefficient alpha of .96 (Hall & Ponton, 2002).

The data were gathered to understand how students learn mathematics, their perspectives of mathematics throughout the study, and their views of how specific pedagogies influenced their self-efficacy. Each data source (classroom observations, video recordings, interviews, surveys, student journals) was analyzed using standard procedures of coding and/or statistical analysis. The data collection procedures for each source are indicated in Table 1.2.

Table 1.2 – *Data Collection Procedures*

Data Source	Description
Classroom Observations	25 110-minute classroom observations and field notes
Video Recordings	20 hours videotape of students interacting and doing mathematics
Pre/post Interviews	10 15-minute pre-interviews, and 10 20-minute post-interviews
Pre/post Surveys	All 10 participants completed pre/post MSES
Journal Reflections	All 10 participants completed weekly journal reflections

We used three strategies to strengthen the validity and reliability of the study: First, triangulation (Denzin & Lincoln, 1998) was used throughout the analysis process to understand the multiple sources in relation to each other, elucidate clear trends and themes, and fortify the credibility of the study. Second, we performed member checks (Lincoln & Guba, 1985) with participants to ensure accurate transcription and interpretation of the data. Third, we completed inter-rater reliability (Mays & Pope, 1995), where independent coding resulted in 95% reliability among coding between the two researchers. When rare discrepancies did emerge, the researchers discussed the difference in detail, resolved the disparity, and refined the codes and themes.

The data collection for the study is suitable given that our goal is not to generalize to the entire student population (Flyvbjerg, 2006). Instead, our intent is to adequately address the research questions by providing descriptive findings with high validity and reliability. Selecting the aforementioned data sources optimized our opportunities to collect germane data.

We analyzed the data using the constant comparative data analysis method, explained by Glaser and Strauss (1967) as the process by which the researcher shifts back and forth between the data and the field to collect information that will then be coded into themes. As shown in Table 1.3, the data sources were analyzed using appropriate analysis procedures.

Table 1.3 – *Data Analysis Procedures*

Data Source	Data Analysis
Classroom Observations	Multiple coding methods on field notes focusing on peer-to-peer and student-to-teacher interaction and discourse
Video Recordings	Video breakdown and analysis using Studiocode
Pre/post Interviews	Multiple coding methods on student perspectives of pedagogy, learning independently and collaboratively, and self-efficacy
Pre/post Surveys	Various statistical tests using SPSS to determine the effect of the project on enhancing students' self-efficacy
Journal Reflections	Multiple coding methods on student perspectives and perceptions of the learning environment, pedagogy, and class as a whole

Results and Discussion

This study examined a classroom of diverse students learning mathematics through two pedagogical strategies. Because this is a case study, rather than a characterization of a systemic trend, the purpose is to gain rich, detailed data for analyzing student development in this pedagogically driven classroom environment. In this section, we present our findings that pertain to the research questions.

RESEARCH QUESTION 1: What do individual students perceive to be the benefits and challenges of traditional and collaborative pedagogies, particularly with respect to enhancing their mathematics self-efficacy?

To address this first research question, we examined the data using the constant comparative data analysis method. For added validity, we investigated the multiple data sources through triangulation, conducted member checks, and performed inter-rater reliability. After a detailed analysis of the participants' perceptions on the traditional and collaborative pedagogies, we elicited the three most prevalent benefits and challenges of each. In this section, we will discuss the benefits and challenges of traditional instruction and cooperative learning as perceived by the students based on their unique frames of reference.

Before we discuss student perception of the portions of the course that lead to change in their confidence and competency, it is important to establish that student self-efficacy was in fact altered during the course. To this end every student completed the MSES twice, once as a pre-survey taken in the first week of the term, and again as a post-survey administered in the last week of the term. Using SPSS, we conducted descriptive statistics to summarize the data and explore the students' pre/post scores. Figure 1 shows the students' MSES pretest scores, posttest scores, and change in scores (posttest – pretest).

MSES Scores Breakdown

Student	**MSES Pretest Score**	**MSES Posttest Score**	**MSES Change**
Adam	7.18	7.56	+0.38
Angela	7.03	8.21	+1.18
Cassie	2.94	4.76	+1.82
Chelsea	5.74	6.62	+0.88

Dante	6.50	7.00	+0.50
Jason	4.91	6.32	+1.41
Jin	6.76	7.53	+0.77
Juanita	4.13	4.18	+0.05
Rahul	7.41	8.15	+0.74
Sara	5.97	7.29	+1.32

Figure 1 – MSES Pretest, Posttest, and Change in Scores

A more illustrative display of the data distribution can be seen in the clustered bar graph in Figure 2. This graph indicates the students' MSES pretest scores in blue, the posttest scores in green, and the change in scores on the horizontal axis.

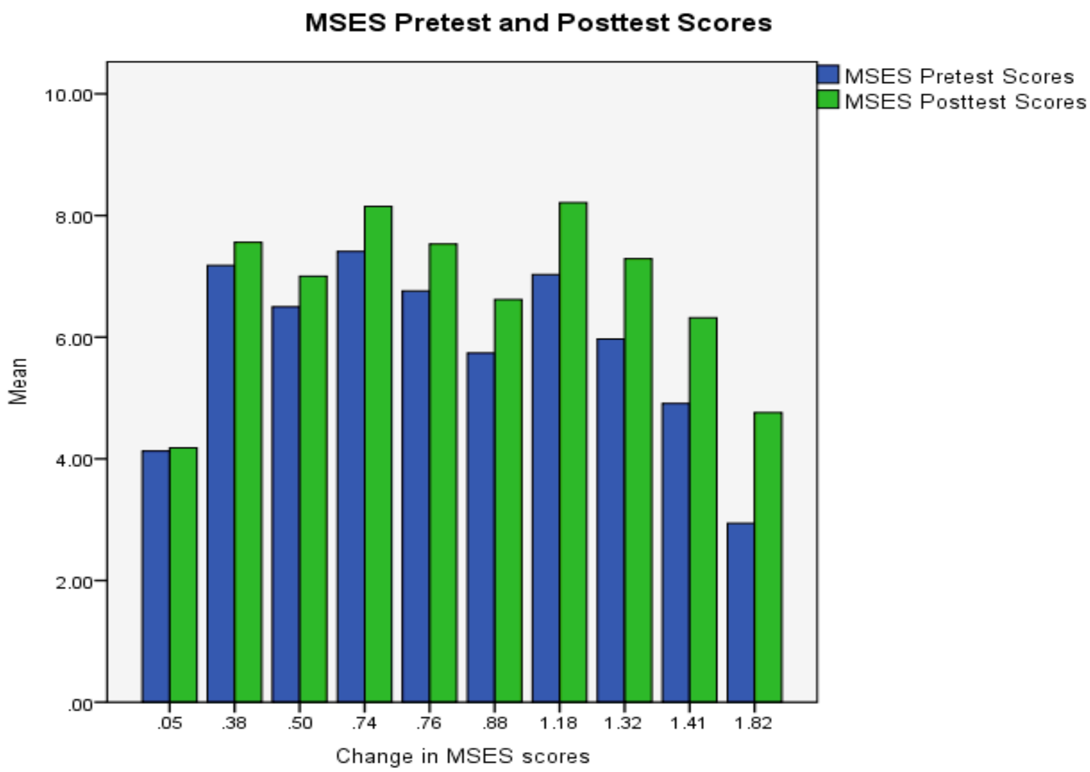


Figure 2 – Clustered bar graph of MSES pretest, posttest, and change in scores

The mean MSES score at the start of the term was 5.86 and the mean score at the end of the term was 6.76, a mean increase of 0.90. To compare the pretest scores with the posttest scores, we used SPSS to perform a paired sample t-test on the data. Our hypothesis was that there is a significant difference between students' MSES pretest scores and posttest scores. No significant difference between students' MSES pretest scores and posttest scores was the null hypothesis for this analysis. Figure 3 displays the results of the paired sample t-test:

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 MSES Pretest Scores - MSES Posttest Scores	-.90500	.53301	.16855	-1.28629	-.52371	-5.369	9	.000

Figure 3 – Paired sample t-test results

The p-value = 0.000 is less than p = 0.05, so, the null hypothesis is rejected, and there is a significant difference between the MSES pretest scores and the posttest scores. Therefore, the results of the paired sample t-test indicate that the students' mathematics self-efficacy increased significantly from the start of the term to the end. With an initial verification from this quantitative data that there existed a significant change in each students' self-efficacy, we return to the student assessment of pedagogical methods that lead to the change.

THREE BENEFITS OF TRADITIONAL INSTRUCTION

Many students enter college with a variety of goals; in order to accomplish these goals and flourish intellectually, students must be intrinsically motivated (Deci, Vallerand, Pelletier, & Ryan, 1991). In the learning process, self-motivation is important to continuously develop one's academic skills while growing introspectively and individually. Given the role of self-motivation and desire to establish oneself as an independent thinker able to succeed without communal scaffolding, the three most prevalent benefits of traditional instruction as perceived by the students are the following:

1) Easier to focus when working independently
2) Convenience of being able to work at one's own pace
3) Increase self-efficacy by solving problems correctly without aid

Easier to focus when working independently

While the professor lectured at the board, the students attentively took notes, rarely chatted, and focused on the task at hand. In the interviews and journal reflections, many students stated that they were less distracted during traditional instruction because they did not have to worry about others, only themselves. In his journal reflection, Dante recorded, "I obtain more when I work at it myself instead of trying to understand where everyone in the group is coming from." Several students expressed feelings similar to Dante's by acknowledging that not having friends to chat with in groups made it much easier to focus. During the interview, Cassie said traditional instruction helped her concentrate more on the task at hand without having the temptation to talk with friends, and Jason described his independent working environment as being much more peaceful and quiet than small groups. The video data substantiated the students' claims that traditional instruction helped them focus more on mathematics. They made constant eye contact with the teacher, concentrated on his lesson and his writing on the board, and compiled detailed notes, which they showed during the post-interviews.

Convenience of being able to work at one's own pace

During traditional instruction, the majority of the students emphasized that an advantage of working independently was the convenience of working at their own pace (i.e., students who preferred to work quickly did not have to slow down to teach their classmates, or other students who did not have to deal with the pressure of keeping up with their classmates). While acknowledging researchers (Blumenfeld, Marx, Soloway, & Krajcik, 1996) who stressed the benefit of students cooperating with each other to overcome such difficulties, this advantage comes straight from the students' perspective. Chelsea stated, "I enjoy working individually more because you can actually see how much you know or have learned. You can work at your own pace." During his interview, Adam explained how traditional instruction allows students to determine how well they comprehend the material with having the option of either working independently or asking the teacher for assistance. This affords students to control their level of intrapsychological learning.

Increase self-efficacy by solving problems correctly without aid

The participants emphasized that the social nature of learning makes working independently more challenging than completing tasks in collaboration with others. Being at a loss when faced with a figurative roadblock can be frustrating and time consuming without the help of a classmate. However, some students felt that working through a perplexing activity can be an effective way to learn and grow. Rahul and Juanita stressed that working individually forces you to teach yourself and understand the material because you are unable to rely on a group mate. Jason and Cassie explicated that although independent work challenges their knowledge, passing the challenge boosts not only their comprehension, but also their mathematics self-efficacy. The danger here is this sense of accomplishment may be fleeting dependent on continued success, but cannot be overlooked as a mechanism for development, as many students shared the same experience. Dante said that traditional instruction "is good for the student to learn how to do things for themselves. Working through problems yourself is probably the best way to learn." He explained that during group work, it is not uncommon for students to rely on others. But, during traditional instruction, it is all on you to figure it out.

THREE CHALLENGES OF TRADITIONAL INSTRUCTION

Prior to the early 1980s, traditional instruction was the pervasive pedagogy of mathematics education. However, as the reform movement began, many researchers criticized traditional instruction for restricting students to merely following teachers' directions, rather than thinking critically and understanding conceptually. Many educators adopted the sociocultural belief that students could learn more through interacting with classmates (discussing and doing), rather than what was seen as passive assimilation of knowledge. Critics of traditional teaching implicated that limiting student interaction limited their growth and development, which ultimately inhibited students in cultivating a richer, deeper view of mathematics (Schoenfeld, 2004). The participants indicated that the three most dominant challenges of traditional instruction are the following:

1) No one to turn to when help is needed
2) Unable to learn multiple viewpoints from classmates
3) Difficulty understanding teacher led lectures

No one to turn to when help is needed

There are myriad reasons why Mathematics 111 (College Algebra) is the most failed course in the State university system. Among the competency related reasons for this statistic is students' inability to contend with mental blocks while problem solving. During traditional instruction, students often become extremely frustrated when they are left to fend for themselves when faced with difficulty moving beyond a convoluted portion of an activity. For example, when asked about her perceptions of traditional instruction, Juanita replied that it is occasionally difficult "because if I don't know how to solve the problem I just sit there without knowing how to solve it. At least with collaborative I have others ideas helping me solve the problem." In response to the same question Angela responded, "you have to think about the problem on your own, and there is not anyone who can do the work for you." In this sense, investigating problems independently without the aid of others challenges students to think critically, which results in some students capitulating and failing to complete the work.

Unable to learn multiple viewpoints from classmates

Learning multiple viewpoints or perspectives is tremendously helpful as it provides students the option to transition from one method to another until an effective analysis technique is realized. Working with classmates during cooperative learning provides students with many opportunities to share and learn multiple solution methods. Working independently does not afford students opportunities to be introduced to ways of thinking and knowing that are different from their own. The majority of the students to varying degrees perceived this lack of exposure to new forms of logic as being a weakness. Sara voiced the limitation by stating, "you don't have the option of asking others for help or learning different methods." Rahul had a similar assessment offering that without having the option of learning different methods, his mathematical knowledge is limited to just what he knows.

Difficulty understanding teacher led lectures

Acknowledging that every teacher is different – some are fantastic lecturers and others are inadequate, the students praised the professor on his ability to teach clearly and effectively. Dante complimented the instructor by saying, "The teacher in this instance was really good and I could really understand what he was trying to teach. So, when he taught I listened and most of it made a lot of sense." Jason and Adam described the professor as having great teaching strategies and that he surpassed expectations by being patient, supportive, and available to his students. Cassie added, "He's willing to help you. He's very nice. The way he just does his lectures makes you pay attention versus being just very monotone. He's very excited about it, I guess, versus a teacher who just reads off a slide." The classroom culture created by such an educator portends that students will be able to have increased teacher-student interaction to prepare them for the intrapsychological learning they engage in both during and subsequent to lecture. Professors who recognize the social and situated nature of learning can make traditional teaching more socially likely through engaging presentation styles, culturally competent exchanges, and an overall dialogue-welcoming demeanor. The video recordings showcased the professor as being a gregarious teacher, regularly teaching with a smile. Cassie asserted that the teachers' affable personality motivated her to pay attention, which helped her to further grasp the material.

THREE BENEFITS OF COOPERATIVE LEARNING

Sociocultural theorists have praised group work for cultivating human development and establishing supportive learning environments (Vygotsky, 1978). Mathematics educators point out that interacting with classmates during group work enriches versatility in problem solving and strengthens comprehension. According to the students' perceptions, the three most prevalent benefits of collaborative learning are the following:

1) Successfully solve problems: Enhance self-efficacy
2) Discuss multiple methods, gain multiple perspectives
3) Increase social capital for future implications

Successfully solve problems: Enhance self-efficacy

In contrast to the prevailing disadvantage of isolated problem solving faced by students in traditional instruction, a significant advantage of cooperative learning is strength in numbers. The sociocultural benefits of clearing up confusion with the help of a classmate should not be minimized. Through a sociocultural lens, group work plays a key role in optimizing interaction, which provides students an elevated state of learning while they investigate and solve problems. Cassie said, "While I'm trying to learn a topic I would rather work in a group to ask questions," and if she does get stuck, then "it also guarantees that I get help if I do not understand something." When asked which pedagogy she prefers, Juanita answered, "Collaborative. Because if I didn't know the answer to something, somebody else like knew it, then we kind of just bounced ideas off each other." Bouncing ideas off of classmates became a common theme with students benefitting collectively from combining their mathematics knowledge.

Discuss multiple methods, gain multiple perspectives

Collaborating with classmates provides students opportunities to learn multiple solution methods, and consequently, gain multiple perspectives of how to analyze and solve problems. Students claimed discussing multiple methods and gaining multiple perspectives helped them become more versatile, and if they ever got stumped using one method, then they could always switch to another. Adam said he likes cooperative learning because "It's nice to see other people's views and approaches to solving problems," and "when we are forced to break into groups, and study together, it allows multiple viewpoints to merge into one cohesive group." Adam asserted that he enjoyed learning multiple viewpoints from his group mates because he was able to become a more proficient and refined student.

Increase social capital for future implications

The previous benefits of cooperative learning occurred in the real-time; however, group work also has forthcoming effects, as students built lasting relationships and increased their social capital for future coursework. Working collaboratively provides students with opportunities to practice team-building strategies and gain an understanding of team dynamics so they can be better prepared to work in teams in the future. This is a significant development given the necessity of group study to meet learning expectations in upper level STEM courses. While studying diverse Calculus students, Treisman (1992) discovered that Asian students regularly created fruitful partnerships adept at productively exercising team-building skills. These students not only learned how to create effective teams, but also how to thrive in groups, a strategy they could parlay into future success in both secondary education and their careers.

THREE CHALLENGES OF COOPERATIVE LEARNING

Geertz (1973) described sociocultural theory as learning through an ongoing process dictated by the individual interacting with others and the environment. Within social and cultural contexts, thinking and doing become intertwined (Rogoff, 2003; Wertsch, 1991). In order for cooperative learning to optimize growth and development, students must be capable of working harmoniously with their classmates. Sometimes this may be difficult as the students indicated that the three most prevailing challenges of cooperative learning are the following:

1) Unwilling/unmotivated group members
2) Possibility of embarrassment
3) Takes more time and effort

Unwilling/unmotivated group members

The most reported challenge of cooperative learning as perceived by the participants is having an unwilling and/or unmotivated group member. Unfortunately, ever since the inception of cooperative learning, this has been a difficult challenge to tackle (Johnson & Johnson, 1990; Slavin, 1995). Most students asserted that they did not have any problems working with their group members in this class; however, there were isolated reports that working with their classmates could be frustrating. Reasons they provided included group members who relied on others to do the work, solved problems at a slower speed, or people they did not get along with.

Possibility of embarrassment

A common fear of many college students is becoming embarrassed in front of peers. There are myriad of reasons why students may experience this emotion during group work. For example, this is a collaborative environment, yet students are still interacting while harboring the legacy of preexisting insecurities related to mathematics competency. Moreover, the overall context of American culture suggests that we should strive to be the leader. Performing subpar in relation to peers may lead to verification of these insecurities or feelings of public failure. Participants in this study perceived the emotions such instances bring as a weakness of cooperative learning. Juanita said there were moments during group work that "I feel less confident because [my classmates] need to help me get a better understanding and I am just slowing their rolls." Chelsea claimed that one of her "pet-peeves" of group work was slowing down her classmates because of two reasons: 1) She felt like she was holding them back, and 2) She felt pressured to keep up.

Takes more time and effort

Several students felt collaborative learning was challenging because it took more time and effort than traditional instruction. They pointed out that it took longer to discuss problems with classmates, explain the analyses to each other, and ascertain if everyone was on the same page. Sara stated that a hindrance of group work was that "you actually have to discuss what you just learned and show how to do what you learned." Students participating in groups must think not only for themselves, but also for their fellow group members, which caused some students to feel frustrated. Contrary to these students' opinions, researchers indicate that in order to improve achievement rates, it is important for students to experience frustrating moments. Overcoming these moments by successfully solving problems leads to students gaining a sense of accomplishment (Yeung, 2009). Furthermore, the frustration in this instance is due to the challenge of becoming accustomed to group dynamics, a necessary endeavor from a sociocultural standpoint. However, several participants disagreed with the researchers as they attributed adversity to forming a roadblock and making mathematics difficult.

RESEARCH QUESTION 2: What evidence exists indicating the role of discourse in improving diverse students' abilities to learn and do mathematics?

To attend to the second research question, we focused our investigation on the data that highlighted students interacting through discourse, particularly the classroom observations and video recordings (student interactions), journal reflections (students writing about their interactions), and interviews (students speaking about their interactions). In this section, we will present literature on discourse encompassing the benefit and impact it has on student development. Then, through a sociocultural lens, we will discuss evidences of the participants' perceptions of discourse and juxtapose it with the literature.

Discourse promotes students' cognitive development

According to the NCTM (2000), "communication is an essential part of mathematics and mathematics education" (p. 60). Since the NCTM published new material promoting discourse, many mathematics educators have implemented the usage of discourse into the curriculum. Research indicates that learning through discourse is central to helping students develop their mathematical understanding and skills. Discourse has helped students summarize and synthesize mathematics (Chaplin et al., 2009), and improve their cognitive development (Cobb et al., 1993).

In this study, the students' perspectives on discourse aligned with prior research. The participants emphasized that discussing mathematics with their classmates improved their abilities to do mathematics. Through engaging in discourse, students were able to cultivate their mathematical analyzing skills and strengthen their classroom performance. One particular journal reflection question read: "Do you feel discussing math with your classmates helps you?" The students unanimously responded affirmatively providing details showing what they gained. They specified that discussing mathematics collaboratively with their peers refined their skills so that they were able to become more complete students.

Discourse helps increase social capital, which can fortify mathematics comprehension

As the nation's ethnic and racial population continues to rise, diversity in the college student body also increases. In order to recruit and retain minority students, university efforts have grown to include programs to boost academic and social acclimation into the college lifestyle. To help students integrate into college life, strong support systems or social networks are crucial in and out of the classroom. Yosso (2005) defines social capital as "networks of people and community resources. These peer and other social contacts can provide both instrumental and emotional support to navigate through society's institutions" (p. 79). Research studies have indicated that engaging in group work and discussions establishes the classroom as a community and strengthens students' social capital (McKinney, McKinney, Franiuk, & Schweitzer, 2006).

The students' perceptions in this study concurred with prior research studies. As the participants collaborated with each other, they were able to gain interpersonal and mathematics skills, and ultimately enrich their social capital. Through discourse, the data indicated the following three ways students were able to increase their social capital: 1) Students were able to build and strengthen relationships with their classmates, 2) Students (and teacher) were able to establish and maintain a positive, supportive classroom environment, and 3) Looking toward future classes, students were able to build their networks of support.

Discourse plays a key role in enhancing self-efficacy

While considerable research has been performed on studying the effect of group work on student learning, few studies have investigated the influence of student discourse on enhancing self-efficacy. Purzer (2011) explored the relationship between discourse, self-efficacy and achievement by studying 22 first-year engineering students. The participants engaged in discourse regularly, and the study's results indicated that there was a positive correlation between discourse and achievement, as well as discourse and self-efficacy. Hendry and colleagues (2005) examined the effects of collaborative learning on college students' learning styles and self-efficacy. They discovered that through group work, students reported greater self-awareness of their own learning, were more accepting of others' styles, and exited the study with more confidence in their academic abilities.

In this study, 100% of the participants answered a resounding, yes, to one of the final journal reflection questions: "Compared to the start of the term, do you feel your mathematics self-efficacy has increased?” In replying to this question, Adam wrote, “Absolutely. I feel pretty confident in this class… I was able to spend seven, eight hours a week outside of class, at least an hour a day usually more than that.” Adam explained that the main reason his mathematics self-efficacy increased was because of the myriad opportunities he experienced working with his classmates. He pointed out that through discussing mathematics with his classmates he was able to feel more comfortable not just doing it independently, but also explaining it. The video data demonstrated Adam regularly teaching material to his classmates, and he asserted that explaining problems to classmates benefitted him to understand and remember the material more comprehensively. Adam noted that toward the end of the term after reviewing notes, preparing for tests, and working with classmates, “the ideas are just going to fit in your head.”

RESEARCH QUESTION 3: What are the implications of addressing diverse student perspectives when implementing either pedagogy?

The students considered both pedagogical strategies to be beneficial in helping them learn and do mathematics. While both strategies came with limitations, as previously discussed, the students unanimously perceived the pedagogies’ strengths to far outweigh their weaknesses. Dante declared, “I think there are really good things to come out of both strategies,” and when asked which pedagogy is the most effective, Adam stated, “I would say a good combination of both.” Collectively, the students emphasized that the preferred instructional method is a hybrid classroom model, a fine balance between traditional instruction and cooperative learning, where students work independently first, and then in groups to answer questions.

The participants advocated a hybrid model using two unique pedagogies because it attends to the students’ various learning styles. Chelsea explained, “I like a combination of both. They are both helpful in learning. Independently, you can challenge yourself and see what you have learned or know. Collaborative group work helps you to share knowledge and other ways of problem solving.” A hybrid course comprises of students working collaboratively and independently, in and out of groups, through engaging in discourse and working silently. A hybrid model is effective because it is as diverse as the students it serves.

Conclusion

According to the NCTM (2000), qualities of an effective teacher include “knowing and understanding mathematics,” as well as their “students as learners, and… a variety of pedagogical and assessment strategies” (p. 17). This study contributed key components to NCTM’s proclamation by detailing successful attributes of a mathematics professor, pedagogical strategies to enhance classroom learning, and various methods to increase students’ self-efficacy.

The results of this study showed the usefulness of a hybrid classroom model for attending to individual student needs and enhancing their self-efficacy. The outcomes also suggested the hybrid model would be beneficial to introductory courses, where students of varied backgrounds congregate to become proficient with foundational concepts. Moreover, this collection of student perspectives offered evidence that components of a hybrid model can be optimized for full effect on learning outcomes. For diverse STEM students, this setting was a venue in which to strengthen content knowledge while participating in the discourse of the STEM communities to which they strive to gain membership. Increasing diverse students’ self-efficacy in this way can lead to similar future interactions and identification with peers and authorities in STEM fields; an important consequence, as engagement and self-identification with STEM professionals has been reported to lead to an increase in the likelihood of retention (Carlone & Johnson, 2007).

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, *84*, 191–215.

Bauersfeld, H. (1980). Hidden dimensions in the so-called reality of a mathematics classroom. *Educational Studies in Mathematics*, *11*(1), 23–41.

Bettinger, E. P., & Long, B. T. (2006). *Addressing the needs of underprepared students in higher-education: Does college remediation work?* (No. NBER Working Paper No. 11325). Cambridge, MA: National Bureau of Economic Research.

Blumenfeld, P. C., Marx, R. W., Soloway, E., & Krajcik, J. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher*, *25*(8), 37–40. doi:10.2307/1176492

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *The Teachers College Record*, *110*(3), 608–645.

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, *44*, 1187–1218.

Chaplin, S. H., O'Connor, C., & Anderson, N. C. (2009). *Classroom discussions: Using math talk to help students learn* (2nd ed.). Sausalito CA: Math Solutions Publications.

Cobb, P., Wood, T., & Yackel, E. (1993). Discourse, mathematical thinking, and classroom practice. In N. Minick, E. A. Forman, & C. A. Stone (Eds.), *Education and mind: Institutional, social, and developmental processes* (pp. 91–119). New York: Oxford University Press.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, *26*(3-4), 325–346. doi:10.1080/00461520.1991.9653137

Denzin, N. K., & Lincoln, Y. S. (1998). *Strategies of qualitative inquiry*. Thousand Oaks, CA: Sage Publications.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry*, *12*(2), 219–245.

Geertz, C. (1973). *The Interpretation Of Cultures*. New York: Basic Books Classics.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Thousand Oaks, CA: Sage Publications.

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education*, *34*(1), 37–73.

Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education*, *20*, 263–271.

Hall, J. M., & Ponton, M. K. (2002). A comparative analysis of mathematics self-efficacy of developmental and non-developmental freshman mathematics students. Presented at the Mathematics Association of America, Louisiana / Mississippi section.

Hendry, G. D., Heinrich, P., Lyon, P. M., Barratt, A. L., Simpson, J. M., Hyde, S. J., … Mgaieth, S. (2005). Helping students understand their learning styles: Effects on study self-efficacy, preference for group work, and group climate. *Educational Psychology*, *25*(4), 395–407. doi:10.1080/01443410500041706

Huang, G. G. (2000). Mathematics achievement by immigrant children: A comparison of five English-speaking countries. *Education Policy Analysis Archives*, *8*(25), 1–14.

Johnson, D. W., & Johnson, R. T. (1990). Social skills for successful group work. *Educational Leadership*, *47*, 29–32.

Lave, J., & Wenger, E. (1991). Legitimate peripheral participation in communities of practice. In *Situated learning: Legitimate peripheral participation* (pp. 89–117). New York: Cambridge University Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

Mays, N., & Pope, C. (1995). Qualitative research: rigor and qualitative research. *British Medical Journal*, *311*, 109–112.

McKinney, J. P., McKinney, K. G., Franiuk, R., & Schweitzer, J. (2006). The college classroom as a community: Impact on student attitudes and learning. *College Teaching*, *54*(3), 281–284.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *18*, 30–38.

National Center for Education Statistics. (2009). The nation's report card. Math highlights 2009. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from http://nces.ed.gov/nationsreportcard/mathematics/

National Council for Accreditation of Teacher Education. (2008). *Professional standards for the accreditation of teacher preparation institutions*. Washington, DC: Author.

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

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

Purzer, S. (2011). The relationship between team discourse, self-efficacy, and individual achievement: A sequential mixed-methods study. *Journal of Engineering Education*, *100*(4), 655–679. doi:10.1002/j.2168-9830.2011.tb00031.x

Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press.

Schoenfeld, A. H. (2004). The math wars. *Educational Policy*, *18*(1), 253–286.

Slavin, R. E. (1995). *Cooperative Learning: Theory, Research and Practice* (2nd ed.). Boston, MA: Allyn and Bacon.

TIMSS. (2007). Trends in International Mathematics and Science Study [TIMSS]. *Mathematics Achievement*. Retrieved January 11, 2012, from http://nces.ed.gov/timss/results07.asp

Treisman, U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal*, *23*(5), 362–372.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

Yeung, B. (2009). Kids master mathematics when they're challenged but supported. *Edutopia*. Retrieved March 12, 2013, from http://www.edutopia.org/math-underachieving-mathnext-rutgers-newark

Yosso, T. J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education*, *8*(1), 69–91.

IDENTIFYING CHANGE IN SECONDARY MATHEMATICS TEACHERS' PEDAGOGICAL CONTENT KNOWLEDGE

Melissa Goss
U. of Northern Colorado

Robert Powers
U. of Northern Colorado

Shandy Hauk
WestEd

Like several other research groups, we have been investigating measures for capturing change in middle and high school teachers' mathematical pedagogical content knowledge (PCK). This report focuses on 14 teachers who have completed a distance-delivered master's degree in mathematics education. The group is the first of five cohorts who will complete a program that seeks to develop content proficiency, intercultural competence, and pedagogical expertise for teaching mathematics. Analysis included pre- and post-program data from observations of participants at work and written PCK assessments. Results indicate significant changes in curricular content knowledge and discourse knowledge. Path analyses suggest teacher discourse knowledge as measured by the written assessments is significantly related to discourse knowledge as measured by the post-program observation.

Key words: Pedagogical content knowledge, in-service teachers, professional development

Background

In response to the call for advanced professional education accessible to in-service teachers, the Mathematics Teacher Leadership Center (Math TLC), an NSF-funded Mathematics and Science Partnership project has developed and is researching a virtual master's program in mathematics education. The primary goals of the program are to develop content proficiency, cultural competence, and pedagogical expertise for teaching secondary grades mathematics (grades 6 to 12). To document the development of mathematics teaching expertise, project research investigates the pedagogical content knowledge of teacher participants before their enrollment into and after their completion of the master's degree program. Earlier reports have offered preliminary results on data collected mid-program. This report is the first to include pre- and post-program data for the first cohort of graduates.

Pedagogical content knowledge (PCK) is a construct described by Shulman (1986) and subsequently refined by other researchers to encompass the unique collection of discipline-connected knowledge needed for teaching. As PCK has become widely utilized in research on teacher development, the idea of "mathematical knowledge for teaching" has emerged as a useful construct (Ball, Hill, & Bass, 2005). In particular, in seeking to capture what elementary grades teachers do in the teaching of mathematics, the focus has been the question: What mathematical reasoning, insight, understanding, and skills are required for a person to teach elementary mathematics? Many have worked to develop measures to address this question, most notably Ball and colleagues (Ball, Thames, & Phelps, 2008; Hill, Ball, & Schilling, 2008). In their work they have defined three types of PCK: knowledge of content and students, knowledge of content and teaching, and knowledge of curriculum. Their perspective is that thinking and decision-making in teaching requires integrating knowledge from each of three mathematically rich contexts (content and curriculum, content and teaching, content and students).

Many challenges in measuring PCK when it is framed in this way have been reported (Hill, Ball, & Schilling, 2008). Most test development is for K-8 teaching includes just a bit of algebra and little in the way of advanced mathematics and its syntax, such as are found in college mathematics. For the purposes of this research, we use an expanded model of PCK. Based on the work of Ball and colleagues, it includes attention to the mathematical communication that emerges in advanced mathematical thinking, including algebra and proof-based mathematics.

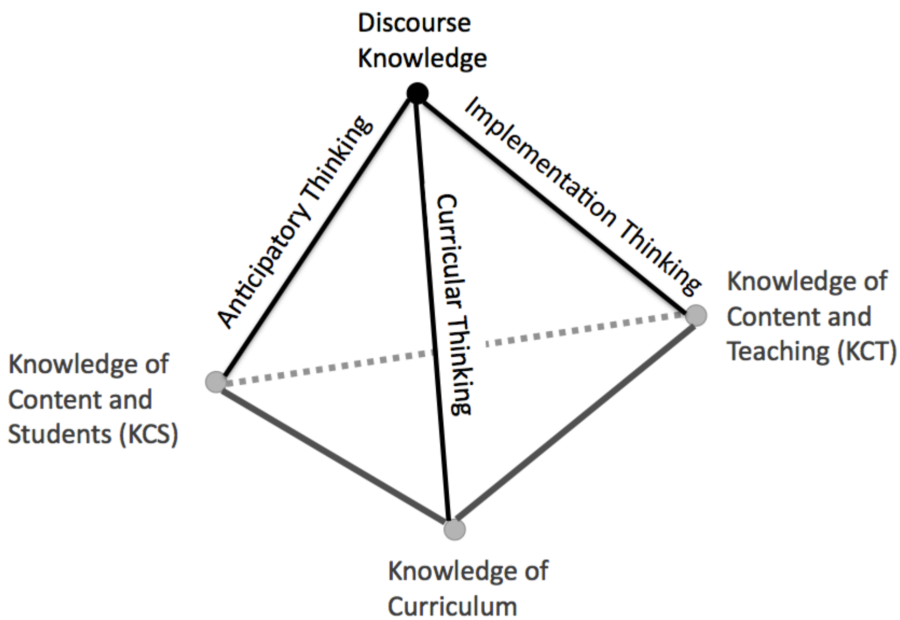


Figure 1. Tetrahedron to visualize relationship among PCK model components. Corners of the base are the aspects of PCK articulated in Hill, Ball, and Schilling (2008).

Working from the foundational three components proposed by Ball et al., the model adds a fourth node of knowledge needed for teaching, discourse knowledge (this aspect brings to the modeling of PCK the mathematical syntax that was part of Shulman's (1986) original description). One way of visualizing the model is as a tetrahedron whose base is the model of Ball et al., with apex of discourse knowledge (see Figure 1). In building this extended theory, our attention has focused on discourse knowledge and the three "edges" connecting it to the components in the Ball et al. model (Hauk, Jackson, & Noblet, 2010; Hauk, Toney, Jackson, Nair, & Tsay, 2013). To situate the results, these four aspects are summarized here.

Discourse knowledge (**DK**) is knowledge about the culturally embedded nature of inquiry and forms of communication in mathematics (both in and out of educational settings). This collection of ways of knowing includes syntactic knowledge, "knowledge of how to conduct inquiry in the discipline" (Grossman, Wilson, & Shulman, 1989, p. 29).

Curricular Content Thinking (**CCT**) is the strategies and approaches to using substantive knowledge about topics, procedures, and concepts along with a comprehension of the relationships among them and conventions for reading, writing, and speaking them in school curricula. In its most robust form, this part of PCK contributes to what Ma (1999) called "profound understanding of mathematics" (p. 120).

Anticipatory Thinking (**AT**) is noticing and strategizing about the diverse ways in which learners may engage with content, processes, and concepts. Part of anticipatory growth involves "decentering" – building skill in shifting from an ego-centric to an ego-relative view for seeing and communicating about a mathematical idea or way of thinking from the perspective of another (e.g., eliciting, noticing, and responding to student thinking).

Implementation Thinking (**IT**) is ways of thinking about how to enact in the classroom the decisions informed by knowledge of content and teaching along with discourse understandings. This includes adaptive, in-the-moment, shifting according to curricular and socio-cultural contexts.

This paper describes the research team's efforts to gather evidence of PCK based on this four-part framework. Data came through written assessments and classroom observations of practice. We report here on our progress to date in addressing the following research questions:

(1) Do teacher-participants differ on PCK as measured by the observation instrument pre-program to post-program?
(2) Do teacher-participants differ on PCK as measured by a written assessment pre-program to post-program?
(3) What is the nature of the relationship between pedagogical content knowledge as demonstrated theoretically on the written assessment and in practice on the observation instrument?

We readily acknowledge the limitations of this non-experimental, small-*n* study. As part of the larger body of work to build theory and associated measurements for PCK, the work reported here is valuable in building a foundation. Note that the intention is not to make causal claims. Rather, we are in the early work of testing predictive validity for instruments and exploring potential avenues for capturing PCK and documenting its change.

Methods

Setting and Participants

The setting was a blended face-to-face and online delivered master's degree program in mathematics for in-service secondary teachers. Designed to reach urban, suburban, and isolated teachers in rural areas, the program is conducted using a variety of technologies (e.g., *Collaborate* for synchronous class meetings, *Edmodo* for asynchronous communication). Offered through a joint effort at the Universities of Northern Colorado and Wyoming, cohorts of 16 to 20 new students each year complete a 2-year master's program in mathematics with an emphasis in teaching (about half of course credits in mathematics, half in mathematics education). Cohort 1 participants teach grades 7 through 12 at schools scattered across the two states (see map, right). The program employs both online and hybrid instruction. Teacher-participants take a combination of face-to-face, hybrid two-location site-to-site, as well as synchronous online courses. In the 30 credit-hour program, 18 credits are in the foundations of secondary mathematics (e.g., modern geometry, continuous mathematics), 6 credit-hours are in mathematics education, including a course developed by the Math TLC project called *Culture in the Math Classroom* (Bartell, Novak & Parker, in preparation), and 6 credits are research-focused (a survey of research in mathematics education and an action research thesis project). Of the 16 teachers who started in the first cohort, 14 completed the coursework of the master's program. All 14 completed the pre- and post-program PCK written test. Though we had pre-program observations for all teachers, for this report we had post-observations for 10 of the teachers.

Instruments

The PCK Assessment included released items from the Learning Mathematics for Teaching instrument (LMT; Ball et al., 2008), new items with more complex mathematical ideas modeled on the LMT items and some *Praxis* items. All of these were limited option "multiple choice" items. For some of these limited option items we added open-ended extensions. Multi-year development of the PCK written test included cognitive interviews with in-service teachers and mathematics teacher educators as they completed individual items or several constellations of items (Hauk, Jackson, & Noblet, 2010). The research team created an alignment of the four PCK constructs across items. For example, one item may be identified as presenting both curricular content and discourse knowledge challenges, while another may foreground curricular and anticipatory thinking. These "loadings" of multiple PCK constructs to items was a purposeful part of the non-linear model underpinning the test design. Each item on the written test loaded on at least two of the four PCK constructs. Consequently, factor analysis was not appropriate given this confounding of variables. In addition to the established face validity of the tests, we conducted tests of the constructs' internal consistency (Cronbach's alpha).

The PCK pre-test showed good reliability overall (α=.81), with good reliability on Curricular thinking (α=.81), acceptable reliability on Discourse knowledge (α=.76), and marginal reliability on Anticipatory thinking (α=.55) (George & Mallery, 2003). While the PCK post-test had acceptable reliability overall (α=.75), we did not see at least marginal reliability on the anticipatory thinking (AT) item set. Because of the variable reliability on the anticipatory thinking construct on the written PCK tests, we did not conduct analyses on it. The observation instrument, based on the LMT video observation protocol (see LMT website; development reported elsewhere) showed good reliability overall (α=.85); including good reliability on Curricular content thinking (α=.84), Discourse knowledge (α=.89), and Implementation thinking (α=.85); and acceptable reliability on Anticipatory thinking (α=.78).

Data Collection

The research team conducted pre-program observations in teacher-participant classrooms in the spring semester prior to teachers entering their first course of the master's program. The team conducted post-program observations in the spring semester three years later; this was two semesters after teachers completed the program. The post-observation data are from a year after completion of the program to give teacher practice time to settle (and avoid detection of an implementation effect that may not be sustained). For both sets of observations, the team observed teachers for three consecutive class meetings (the same researcher(s) visited across the meetings). Like the LMT video protocol, the observation tool used interval recording (Gall, Gall, & Borg, 2007). Interval segments were 6 minutes each: 3 minutes observed, 3 minutes to record the observation by identifying the presence or absence of each protocol category in the observed segment on the protocol form and to record associated field notes. Each class visit had 7 to 12 segments. Experienced observers trained new observers to use the instrument; new raters practiced using the protocol on video data, conducted their first observations of teachers in tandem with an experienced observer, and team members met to calibrate ratings and reconcile disagreements. Inter-rater reliabilities were greater than 0.8 at each calibration check.

Teacher-participants completed the written pre-test at the beginning of their first class session in the program. Of the 14 teacher-participants who completed the program, 9 completed the post-program written assessment at the program closure meeting; for the 5 unable to attend the meeting, members of the research team administered the test at the teachers' school of employment. Members of the research team created answer keys for multiple choice items and a scoring rubric for short answer items. The rubrics were informed both by expected or desired responses created by item developers as well as cognitive interview data. The procedure for developing the rubric was (1) write a desired response, (2) list other anticipated responses, (3) read the responses from a subsample of participants, (4) come to consensus on a scoring rubric. Two or more research team members scored tests separately, compared scores, and met to negotiate and reconcile any disagreements.

Data Analysis

To date the research team has observed 10 teachers after completion of the program. The counts for each of the variables were summed and divided by the number of segments observed to report a relative frequency for each variable for each teacher. A teacher having a score of 23.25 on "Explicit Talk about Math" means that the rater(s) observed the teacher exhibit explicit talk about mathematics during 23.25% of the segments observed. On both the written assessments and the observation instrument, researchers calculated percent scores for each construct by summing teacher scores on items coded for the construct and taking the percent out of total points possible on each construct. To answer the research question of the impact of the Math TLC master's program on teachers' pedagogical content knowledge, we compared entrance and exit data from the PCK assessment and the PCK observations using paired-samples *t*-tests.

To model the relationship between teachers' PCK during practice (as measured by the observations) and teachers' PCK during reflection (as measured by the written tests), we conducted a path analysis on the PCK constructs. The model considered the pre-test and pre-observation scores as exogenous variables. Taking the assumption that change in knowledge leads to change in action, the model examined the effects of the exogenous variables (pre-scores) on the written post-test of PCK, then examined the effects of those three variables on the post-observation scores. The path analysis is for Curricular content thinking and Discourse knowledge, since the reliability of the Anticipatory construct was not sufficient and written assessment did not measure Implementation.

Results

PCK Written Assessment

Table 1 presents the results from paired samples *t*-tests on teachers' percent scores on the constructs of Curricular content thinking and Discourse knowledge for the written tests. Scores on items coded as Discourse knowledge (DK) increased significantly (t=2.189, p=.047) from the pre-test (M=56.82, SD=15.43) to the post-test (M=66.22, SD=19.09).

Table 1. Paired Samples t*-tests for PCK test Cohort 1. Values are percentages.*

	Pre-program (*N*=14)		Post-program (*N*=14)			
PCK Construct	*M*	*SD*	*M*	*SD*	*t*	*p*
Curricular content	60.93	14.57	65.48	16.74	1.261	.229
Discourse	56.82	15.43	66.22	19.09	2.189	.047*

Observations

Tables 2 and 3 present the comparisons of pre-program and post-program observations for the 10 teachers for whom complete data are available. Each table presents the means, standard deviations, and the results of a paired samples *t*-test on each observed variable. Table 2 gives differences on the observation categories aggregated into the four PCK constructs. Table 3 unpacks the information in Table 2 and presents the differences between pre-program and post-program on each item that made up the observation instrument. Because of the number of statistical analyses performed, a cutoff p value of 0.0015 (rather than 0.05) is appropriate, based on a Bonferroni correction (Bland & Altman, 1995).

The results in Table 2 indicate increases approaching significance in two constructs: an increase in score for Curricular content thinking (t=4.31, p=.002) from pre-program (M=45.12, SD=13.18) to post-program (M=56.64, SD=1.066); and an increase in Discourse knowledge (t=3.92, p=.004) from pre-program (M=48.25, SD=13.47) to post-program (M=61.27, SD=10.27). Scores in both Implementation and Anticipatory thinking increased, but the difference was not significant at the 0.0015 level.

Table 2. Paired samples t-*tests for PCK Constructs from Observation Instrument*

	Pre-program (N=10)		Post-program (N=10)			
PCK Construct	*M*	*SD*	*M*	*SD*	*t*	*p*
Curricular content	45.12	13.18	56.64	10.66	4.31	.002@
Discourse	48.25	13.47	61.27	10.27	3.92	.004@
Anticipatory	44.18	12.97	54.26	17.56	1.95	.083
Implentation	59.16	15.13	66.12	9.97	1.54	.159

@ approaching significance, $p < .015$; * significant, $p < .0015$

In all tables, all values increased, which may indicate pedagogical effectiveness of the master's program – the program designers had among their goals that the behaviors in the observation protocol would increase in frequency as a result of the program (with the exception of the reverse-worded "Error-Not Present" category). In Table 3, with the adjusted threshold for alpha, there are two statistically significant results. One was in the observation category "General language for expressing mathematical ideas (overall care and precision with language)." While careful use of general language was seen, on average, in about 49% of pre-program classroom segments, by the end of the program it was present in more than 80% (M=80.34, SD=19.71). The other significant result was in the category "Mathematical descriptions (of steps)" (i.e., segments where the teacher or students accurately and clearly described the steps of some mathematical process). On average, across pre-program observations, this was seen in about 40% of class segments (M=40.28, SD=21.94), increasing to almost 70% of the time, post-program (M=68.10, SD=19.31). Three other observed variables approach significance (i.e., p<.01): the percent of segments where (a) student voices were present in the room (increasing from 80% to 90% of segments), (b) teachers were observed to use conventional notation (increasing from 54% to 90% of segments), and (c) fewer mathematical errors occurred (decreasing from about 4% of the time to nearly 0%).

Table 3. Paired Samples t-*tests for Observation Protocol Variables.*

	Pre-program (N=10)		Post-program (N=10)			
Observation Item	*M*	*SD*	*M*	*SD*	*t*	*p*
Format for Segment						
Whole Group	51.61	22.16	63.39	18.40	1.170	.272
Small Group	23.79	22.51	39.26	22.51	2.832	.020
Individual	41.70	26.03	28.98	12.15	-1.610	.142
Lesson/Segment Type						
Review	26.46	18.03	22.46	14.17	-0.549	.596
Introducing tasks	7.23	4.74	10.64	5.31	2.262	.050
Student work time	45.00	24.16	50.04	16.76	0.586	.572
Direct instruction	24.15	15.27	33.00	16.53	1.201	.260
Synthesis or closure	5.77	4.91	8.10	6.02	1.147	.281

[Table 3 continues on the next page]

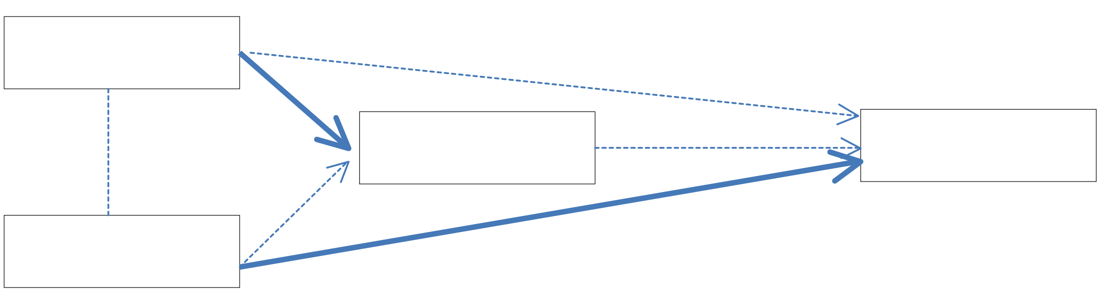
Table 3-Continued. Paired Samples t*-tests for Observation Variables.*

	Pre-program (*N*=10)		Post-program (*N*=10)			
Observation Item	*M*	*SD*	*M*	*SD*	*t*	*p*
Math Teaching Practices						
Voices – Students	79.82	18.32	89.29	15.88	3.375	.008@
Voices – Teacher	80.77	21.98	93.81	8.23	1.949	.083
Real-world Problems	26.55	28.47	36.50	32.44	.826	.430
Interprets Students' Work	63.33	18.39	73.01	13.75	2.120	.063
Explicit about Tasks	82.20	16.20	87.52	12.42	.916	.384
Explicit Talk about Math	59.03	27.18	75.59	11.91	1.801	.105
Explicit Talk about Reasoning	29.93	23.18	49.48	17.94	2.821	.020
Instruction Time	86.10	10.20	87.02	6.81	.249	.809
Encourages Competencies	67.07	26.83	45.04	40.52	-1.420	.189
Knowledge of Math Terrain						
Conventional Notation	54.39	21.38	79.95	15.25	3.353	.008@
Technical Language	72.59	17.38	77.67	13.53	.760	.467
General Language	49.06	13.88	80.34	19.71	4.528	.001*
Selection for Ideas	87.17	8.31	91.16	5.68	1.989	.078
Selection to Represent Ideas	31.70	23.97	43.64	25.51	1.892	.091
Multiple Models	17.80	14.90	33.69	24.20	2.138	.061
Records Work	59.67	28.25	52.01	20.20	-.585	.573
Math Descriptions	40.28	21.94	68.10	19.31	5.003	.001*
Math Explanations	40.65	23.26	55.80	16.29	1.782	.108
Math Justification	14.32	16.13	23.09	11.05	1.928	.086
Math Development	84.50	16.57	88.67	6.11	.753	.471
Errors – Not Present	96.27	2.67	99.78	0.69	3.858	.004@

@ $p < .015$, * $p < .0015$

Relationship between Observation and Written Assessment

The figures below display the results of path analyses exploring the relationships among the program's pre-scores and post-assessment and post-observation scores. Figure 2 shows the full model for Curricular content thinking (CCT). There was a significant effect of the Pre-Test (β=.88, *SE*=.17, $p < .01$) and no significant effect of the Pre-Observation on the Post-Test of CCT. There was a significant effect of the Pre-Observation (β=.68, *SE*=.17, $p < .05$) on the CCT Post-Observation. There was no significant effect of the Pre-Test on the CCT Post-Observation, but interestingly the effect was negative (β=-.60, *SE*=.27). Although the effect of the Post-Test on the Post-Observation was relatively high (β=.81, *SE*=.28), it was not significant, which may be due to the small *n* and large standard error.



ure 2. Path diagram for CCT construct.

Figure 3 shows the full model for Discourse knowledge (DK). There was a significant effect of the Pre-Test (β=.78, *SE*=.26, *p*<.05) and no significant effect of the Pre-Observation on the Discourse Post-Test score. There was no significant effect of either the Pre-Observation or the Pre-Test on the Post-Observation of DK, although, like CCT, the effect of the Pre-Test was negative (β=-.58, *SE*=.19). Finally, there was a significant effect of the Post-Test on the DK Post-Observation (β=.92, *SE*=.17, *p* < .05).

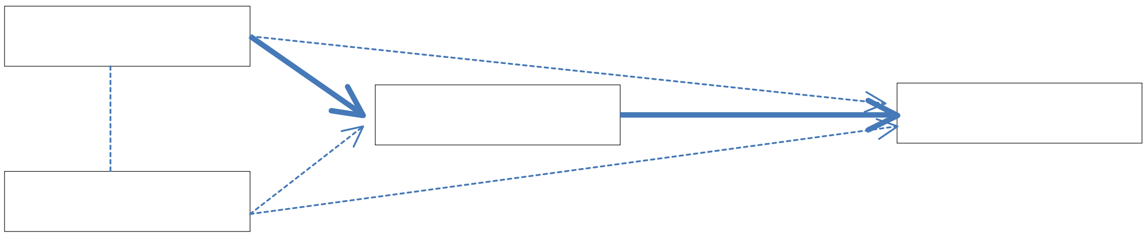


Figure 3. Path diagram for DK Construct.

Discussion

Because of the small sample size, the study is underpowered for full validation of the assessment and observation scores. Additionally, the small sample size makes generalizing the results problematic. What is apparent is that pre- to post-program written test score changes suggested positive outcomes of the master's program in the target area of investigation: development of pedagogical expertise for teaching secondary mathematics, particularly in the communication skills of responsive classroom discourse. The significant increase in curricular content thinking (CCT) from pre- to post-program teacher observations may reflect the master's program emphases on increasing participant understanding of advanced mathematics and deepening secondary-school-level-appropriate conceptual connections. This is evident in some of the significant increases on individual variables in observation categories. For example, the significant increase in the use of conventional notation may indicate that the master's program supported teacher-participants in the habit of using conventional notation to communicate. In addition, the mathematics courses required participants to be explicit about their thinking, reasoning, and justification of answers, which may help explain the significant increase in mathematical descriptions category. However, there were only small (non-significant) increases in the mathematical explanations or the mathematical justification of the reasoning process, so more work needs to be done in the program to support teachers' attention in these important realms of mathematics teaching and learning (perhaps as they challenge the prescribed curricula, which tend *not* to foreground these things). Finally, the reduction in observed errors may indicate a stronger content knowledge for teaching secondary mathematics.

The significant increase in Discourse knowledge (DK) on the written test and in observations may be related to the master's program mathematics education courses. In particular, these courses employed and examined reports of responsive mathematics pedagogy. That is, they made explicit use of the research base on student-centered classrooms and the mathematics instructional practices that support students in the construction of knowledge (rather than the transmission of knowledge by teachers). For example, observers saw significant increases in the percent of small group work and in students' voices in the classroom. This may indicate that teachers' practices shifted to more decentered (or some forms of "learner-centered") approaches. Additionally, the program included several credit hours of reading and writing about mathematics education research focused on the NCTM process standards. The increase in discourse knowledge in general may be attributed partly to the pedagogy courses that allowed participants to read research and experience what good mathematics discourse "sounds like and feels like" (Cohort 1 participant, personal interview, October 8, 2012). Finally, the *Culture in the Math Classroom* course experience included several culture and discourse-specific awareness building activities, scaffolding teachers in decentering their instruction. We suspect that this course and the potential shift in perspectives that teachers may have gained from it will turn out to be a significant predictor of change in Discourse knowledge in the program – part of our ongoing work investigates this hypothesis.

The path analyses relating PCK as demonstrated on the written test and in practice provide interesting results that need further investigation. As noted, the path diagram for Discourse knowledge, Figure 2, suggests that the written test may have predictive value in capturing classroom practice. If this turns out to be a robust result, across populations of teachers, it could reduce or eliminate the need for expensive classroom visits when attempting to determine impact on practice (e.g., it may be the case that pre-program observations and pre/post written tests can give sufficient impact information without the need to re-visit classrooms post-program). Researchers need to conduct further investigation into the ways to measure these constructs and to extend the research to larger, more generalizable samples to verify these results. Additionally, with further data we hope to clarify the degree to which the negative, albeit not significant, correlation of the pre-test with the post-observation is an artifact of small n and within-sample variability or may be an indicator of another variable, possibly related to intercultural orientation and the relative impact of the *Culture in the Math Classroom* course on that orientation.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. DUE0832026. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Ball, D.L., Hill, H., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Mathematical Educator,* 29(1), p. 14-46.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59,* 389-407.

Bartell, T., Novak, J., & Parker, F. (in preparation). Culture in the Math Classroom: A path to improving student success. See preliminary work from a 2011 presentation documented at: http://hub.mspnet.org/index.cfm/msp_conf_2011_abstract2

Bland, J. M., Altman D. G. (1995). Multiple significance tests: The Bonferroni method. *BMJ* [Note: Formerly *British Medical Journal*] 310(6973),170.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007) Educational research: An introduction. Boston: Pearson Education.

George, D. & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference.*11.0 update (4th ed.). Boston: Allyn & Bacon.

Grossman, P. L., Wilson, S. M., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. In M. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 23-36). UK: Pergamon.

Hauk, S., Jackson, B., & Noblet, K. (2010). No teacher left behind: Assessment of secondary mathematics teachers' pedagogical content knowledge. *Proceedings for the 13th*

Conference on Research in Undergraduate Mathematics Education (electronic). PDF available at sigmaa.maa.org/rume/crume2010/Archive/HaukNTLB2010_LONG.pdf

Hauk, S., Toney, A., Jackson, B., & Nair, R., Tsay, J.-J. (2013). Illustrating a theory of pedagogical content knowledge for secondary and post-secondary mathematics instruction. *Proceedings for the 16th Conference on Research in Undergraduate Mathematics Education (electronic).*

Hill, H., Ball, D., & Shilling, S. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372-400.

Learning Mathematics for Teaching (LMT) Website: http://sitemaker.umich.edu/lmt/home

Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahway, NJ: Erlbaum.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

A FORAY INTO DESCRIBING MATHEMATICS MAJORS' SELF-INQUIRY DURING PROBLEM SOLVING

Todd A. Grundmeier
Dylan Retsek
Dara Stepanek

Cal Poly, San Luis Obispo

Self-inquiry is the process of posing questions to oneself while solving a problem. The self-inquiry of thirteen undergraduate mathematics majors was explored via structured interviews requiring the solution of both mathematical and non-mathematical problems. The students were asked to verbalize any thought or question that arose while they attempted to solve a mathematical problem and its nonmathematical logical equivalent. The thirteen students were volunteers who had each taken at least four upper division proof-based mathematics courses. Using transcripts of the interviews, a coding scheme for questions posed was developed and all questions were coded. Data analysis suggests that the "good" mathematics students focus more questions on legitimizing their work and fewer questions on specification of the problem-solving task. Additionally, questions have arisen about the further exploration of self-inquiry.

Key-Words: Problem Solving, Proof, Self-Inquiry, Logic, Questioning

Introduction

Experts in mathematics do things differently than the masses. It therefore makes sense to rigorously study exactly what characterizes expert mathematical thought, ultimately aiming to transfer this understanding to better educate undergraduates in mathematics. Indeed, much recent work in undergraduate mathematics education has explored this very idea. From how experts read proofs (Inglis & Alcock, 2012) and vet the work of their peers (Inglis et al, 2013) to how they make conjectures (Belnap & Parrott, 2013) and use metaphors/perceptuo-motor activity (Soto-Johnson & Oehrtman, 2012), a clearer picture of expert mathematical practices is beginning to emerge.

Many of us who teach mathematics have shared the experience of guiding a student to some understanding through a serious of well-chosen questions, rather than well-chosen answers. In fact, the student often already knows the answers; thus, it is in fact the questions that they are really after. It is therefore natural to wonder what questions (if any) the student asked *her/himself* during their initial consideration of the problem.

S*elf-inquiry* is the process of posing questions to oneself while solving a problem. This paper represents a preliminary foray into the question of expertise through the lens of self-inquiry. Here, though, the expert/novice distinction is made within the category of undergraduate mathematics students, as opposed to comparing the practices of students to those of seasoned mathematicians.

The decision to focus our attention on undergraduate self-inquiry has clear motivation. Perhaps we can isolate the questioning practices of students who are having mathematical success (the "experts") from those of their less successful peers. If so, we can then make more informed pedagogical decisions designed to foster such self-inquiry among all of our students, thereby improving undergraduate mathematics education.

Methodology

While related research has been conducted in secondary education and reading comprehension (Kramarski & Dudai, 2009; King, 1989) and in general mathematical thinking (Schoenfeld, 1992), it seems that the self-inquiry of undergraduates in the process of mathematical problem solving has not been explored. In order to begin this line of inquiry, thirteen mathematics majors were interviewed with the instruction to verbalize every thought process and question that arose as they attempted to solve two problems. The first problem was set in a mathematical context with contrived terminology that was new to all students:

Mathematical Context (Pleasant Sets): A set S of real numbers is called ***pleasant*** if each element of S has both a unique immediate successor in S and a unique immediate predecessor in S.

Let S be a pleasant set. Suppose the numbers a, b, c, d, and e belong to S and satisfy

(i) b is greater than or equal to the successor of d and less than or equal to the predecessor of e;
(ii) a is the successor of d;
(iii) c is greater than or equal to the successor of e.

Put $a, b, c, d,$ and e in numerical order.

The second problem, though posed in a nonmathematical context, was logically equivalent to the first:

Nonmathematical Context (Feeding Time): Zookeeper Jane feeds the animals in 15 minute intervals every morning such that

(i) the giraffes are fed after the monkeys but before the zebras;
(ii) the bears are fed 15 minutes after the monkeys;
(iii) the lions are fed after the zebras.

What is the feeding schedule?

Indeed, the two problems really are isomorphic in the structure of their solutions. In both problems, (i) yields

$$d < b < e \text{ (monkey < giraffe < zebras),}$$

then (ii) yields

$$d < a < b < e \text{ (monkeys < bears < giraffes < zebras),}$$

and finally (iii) yields

$$d < a < b < e < c \text{ (monkeys < bears < giraffes < zebras < lions).}$$

Despite the clear similarities, we will see that the two problems generated quite different question profiles among the thirteen participants and that the difference between high achieving students and their peers was heightened by the pleasant sets problem and diminished by the feeding time problem.

Initial Coding

Upon completion of the thirteen interviews, the three authors independently analyzed each transcript with the aim of identifying all participant generated questions while solving the two problems. After analyzing the transcripts individually the authors reconvened and came to a consensus on the identified questions.

Coming to this consensus was nontrivial given the wide variety of questions in the transcripts. Some questions were obvious and easy to agree upon; for instance, the question "Does successor mean before or after?" is hard to miss. There were, however, more nuanced examples that forced the authors to consider what, exactly, constitutes a question in self-inquiry.

For example, it is not clear whether the absence of a question mark implies the absence of a question. Should the statement "Hmm, let me see if that makes sense" be labeled as a question in self-inquiry? If one interprets it as two sentences

"Does that make sense? I'm going to take a moment to see."

then clearly a question has been asked. On the other hand, if one interprets it as one sentence

"I'm going to see if what I just wrote makes sense."

then perhaps *whether* the student's prior work makes sense is not being questioned. Maybe they are simply checking their work as a matter of course and following the interviewer's request that they verbalize what they are doing.

In the end the authors concurred that "Hmm, let me see if that makes sense" should indeed be categorized as a question. This example is illustrative of the subtlety involved with simply identifying self-inquiry, the classification in this case hinging on that sole "Hmm".

Classifying the questions in each student's transcript allowed the authors to compare both volume and frequency of questions from student to student. Of course, this comparison neglects the differences in *types* of questions. Surely a few well-formulated, incisive questions would trump a whole host of surface-level, off-topic questions? To be certain, the authors required a finer analysis. Thus was born the "Question Tree."

The Question Tree

Question classification begins early in life. Schoolchildren are taught to recognize whether a question falls in the categories who, what, where, when, why or how. Beyond this primitive grouping, however, there is a general dearth of scholarship on categorizing questions. Some related research has been conducted both in secondary education and reading comprehension (Kramarski & Dudai, 2009; King, 1989). Computational linguists have an interest (Moschitti et al, 2007), as do educational researchers (Dillon, 1984; Smith & Mukherjee, 1994).

In the realm of general mathematical thinking we have Schoenfeld's seminal work (Schoenfeld, 1992), and for problem-solving techniques (including suggestions for productive, specific self-inquiry) we have Polya's classic (Polya, 1946). But there seems to be little or no research directly related to classifying self-inquiry.

In order to better differentiate between question types, the authors developed an instrument called the *question tree*. Using Polya's suggested self-inquiry as a springboard, the tree grew from the authors' reflection on their own expert mathematical inquiry and the (sometimes) inexpert inquiry of their own students. The result is intended as a comprehensive tool that may be applied to any question that might occur in the course of mathematical problem solving.

At its broadest, the tree groups questions into one of three main types. *Definition* questions are questions about terms in the problem statement or otherwise relevant to the setting. *Specification* questions are questions designed to elicit more information. Finally, *legitimacy* questions are questions concerning the validity or utility of some action taken by the problem solver.

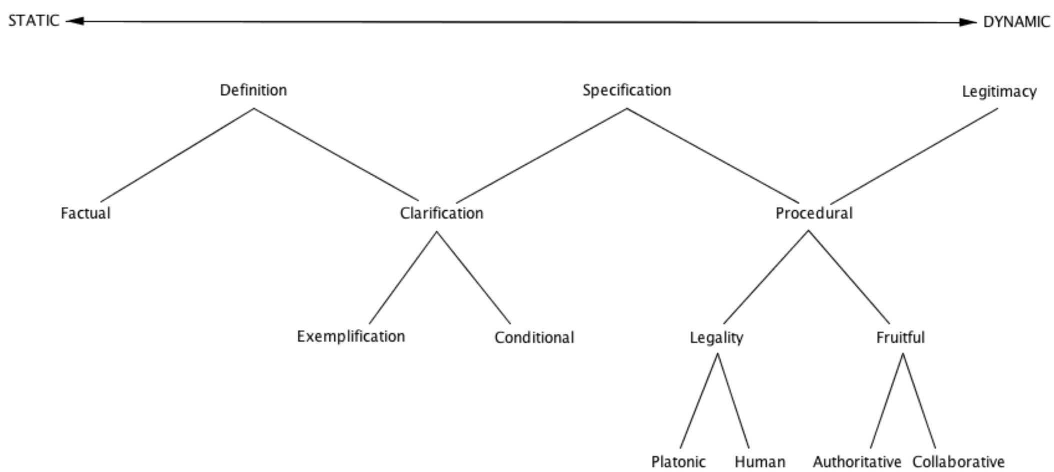


Figure 1: The Question Tree

Within the Definition category, there are two sub-types: *Factual* and *Clarification*, which represent the distinction between denotation and connotation. For example, the question "What is the limit definition of the derivative?" is a factual question about the word 'derivative', while "What does the derivative represent?" is a clarification question about what the derivative *means*.

The Specification category likewise splits into two deeper subcategories: *Clarification* and *Procedural*. Here, clarification elicits more information "Are we in an integral domain here?" whereas "How do I factor a cubic" is a procedural question that asks how something is *done*.

By it's very nature, the third main category deals only in actions. Legitimacy questions ask not *how* something is done, but *whether* it should be (or should have been) done: "I apply the quadratic equation now, right?" Consequently, Procedural questions are a shared root of both Specification and Legitimacy types.

In summary, the second tier of the question tree also has three categories: Factual, Clarification, and Procedural. Factual definition questions bifurcate no further, but Clarification and Procedural questions do bear deeper, more nuanced tendrils.

To wit, Clarification questions fall into two new subtypes: *Exemplification* and *Conditional*. We often use examples to clarify a definition or seek more information. For instance, "Is the dihedral group commutative?" might shed light on the meaning of the word 'commutative' through exemplification. On the other hand, we sometimes gain clarity by adding or removing conditions or hypotheses: "What if the function *f* is merely continuous?"

Similarly, Procedural questions also split into two new subtypes: *Legality* and *Fruitful*, which essentially parse "*can* we" from "*should* we". For example, "Can I cancel here?" is a question of legality of some procedure (compare with the earlier question "Are we in an integral domain here?" which seeks information that could legitimize canceling, whereas the former asks directly for the legitimacy.) Questions of fruitfulness, in contrast, ask whether some procedure was (or would be) a good idea: "Should we group like terms here?"

The final and subtlest tier of the tree further categorizes Legality and Fruitful questions. Within the classification Legality, the subtypes *Platonic* and *Human* aim to distinguish between "Is this legal: $(a+b)/c = a/c + b/c$?" and "Do I have to rationalize this expression in my answer?" The first is a platonic question of correct mathematics within a given context, the immutable rules of the game, while the second is a subjective question of acceptability –posed to a human - rather than correctness.

Likewise, Fruitful questions split into the types *Collaborative* and *Authoritative*, which differentiate between exploration and sanction. The question "Could we try modding out by five here?" is exploratory in nature (collaborating with oneself or some implied observer), while "Should I try modding out by five here?" asks some authority (implied observer?) to sanction the exploration in advance. The difference is subtle, but important. "What if we try to land on the moon?" is bold and full of possibility but "Should we try to land on the moon" is timid and fraught with uncertainty.

Deeper Coding: Implementing the Tree

With the question tree in hand, the authors individually coded the previously agreed upon questions according to their position in the tree. We then reconvened and attempted to reach consensus on the codes. As before, some question codes were easily agreed upon while others generated some debate.

In the end, agreement was reached and every question from every interview transcript was coded. As an illustrative sample, consider the following list of actual interview questions and their assigned codes:

- "Does successor mean after or before? " was coded FD for Factual-Definition.
- "I wonder if I should do that? " was coded CFPL for Collaborative-Fruitful-Procedural-Legitimacy.
- "Let me see if that changes my reasoning. " was coded CCS for Conditional-Clarifying-Specification. [Note the absence of a question mark; statement was interpreted as "Does this new information affect my reasoning?"]
- "Does that work? " was coded PLPL for Platonic-Legality-Procedural-Legitimacy.

There are two final points concerning the question tree. First, the tree seemed to be a reasonably effective coding instrument in that every question fell naturally into some position on the tree. Though it is true that we did not assign every code the tree could offer, every absent code applies to a natural mathematical question. For example, though we did not assign any Platonic-Legality-Procedural-Specification codes, the question "Is cancelation legal in an integral domain?" would receive the PLPS code. This question seeks general information about the mathematical legality of an abstract procedure in an abstract setting. Contrast with "Was that cancelation legal there?" which would be coded PLPL because it asks whether a specific action was mathematically legitimate. It's the ever so subtle difference between "Can I do a u-turn here?" and "Are u-turns legal at four way stops?"

Second, there is a readily apparent left to right progression within the tree. Positions on the left side of the tree represent static questions ("Where is Paris?") while positions on the right represent dynamic questions ("Should we fly to Paris?", "What should we do when we get to Paris", "My aunt lived in Paris, let's ask her" etc.) In general, static questions are those that allow for little to no generative or creative thought, whereas dynamic questions are those that are full of bold possibilities and new ideas.

Certainly static questions may be necessary to begin to solve a problem, but inquiry confined to the left half of the tree is pedestrian and potentially non-generative. Might we expect our experts to visit all corners of the tree as they solve problems? Do expert questions reside mainly in one tier of the tree? On one side? Do experts conduct an intricate dance of inquiry? Are there loops? Left to right waves?

Preliminary analysis does indeed suggest some clear differences between the question profiles of high-achieving members of our cohort and those of their lower-achieving peers.

Results

It should be noted to start that all participants solved both problems correctly but, as will be shown, the required self-inquiry to arrive at those solutions varied. Since all participants were successful at solving both problems, comparing them by problem solving acumen on the task was not feasible.

In order to explore the self-inquiry of "good" students the authors defined the statistic RSQ (Relative Success Quotient) and calculated an RSQ for all participants. To calculate the RSQ the authors focused on the 11 upper division courses that had been taken by at least 7 of the participants. For each course the average GPA and standard deviation of grades were calculated for the last 5 years of course offerings.

A participant's RSQ was then calculated as the average number of standard deviations their grades were away from the mean for the courses they had completed from the 11 chosen. This RSQ calculation identified three clear groups that will be the focus of the remaining data analysis and results.

The groups are named by achievement relative to the other participants. The High group consists of five participants, while the Middle and Low groups consist of four participants each. Table 1 shows the average RSQ and number of questions per problem for each of the three groups.

Table 1: RSQ and Average Questions asked

	Low	**Middle**	**High**
Average RSQ	*-0.167*	*0.479*	*.879*
Average # of Q's Pleasant Sets	13.75	16	9.2
Average # of Q's Feeding Time	2*	3.5	2.8

It should be noted that Kathy was removed (*) from the feeding time analysis because she asked 17 questions in the context of solving the feeding time problem and this was a clear outlier from the other participants. All group Feeding Time data presented herein excludes Kathy, while the Pleasant Set data includes all participants. The data in Table 1 provides an initial glimpse at the self-inquiry of the participants. All groups were able to solve the Feeding Time task by asking a similar number of questions on average. However there seems to be some discrepancy regarding self-inquiry on the Pleasant Sets task.

Self-Inquiry in a Familiar Context

The Feeding Time problem was presented to explore self-inquiry in a familiar context where understanding and contextualizing definitions would not be required. As is highlighted by the data in Table 2, self-inquiry in the context of this problem could be considered fairly uniform amongst the three groups.

Table 2: Self-Inquiry on the Feeding Time Problem by RSQ

	Low	**Middle**	**High**	**Kathy**
Average RSQ	*-0.178*	*0.479*	*.879*	*-.140*
Average # of Q's	2	3.5	2.8	17
Definition Q's (%)	16.7	28.6	14.4	29.4
Specification Q's (%)	0	28.6	42.8	58.8
Legitimacy Q's (%)	83.3	42.8	42.8	11.8

All students quickly completed the problem-solving task except for Kathy and, as indicated previously, her data was not utilized for making comparisons between groups. While the data in Table 2 may imply some differences among the types of questions asked by each group the relatively small numbers of questions makes it difficult to conclude significant differences. The groups all focused more than 70% of their questions on Specification and Legitimacy questions with the Middle group asking the highest percentage and largest number of definition related questions.

Examining the data by question type continues to highlight similarities amongst self-inquiry in this context. For example self-inquiry with respect to Definition questions included all three groups asked CD (Clarifying – Definition) questions at least 75% of the time. For Specification questions, CS (Clarifying – Specification) and PS (Procedural – Specification) accounted for all question types. Similarly for Legitimacy, PLPL (Platonic – Legality – Procedural – Legitimacy) questions account for at least 80% of the questions. It should be noted that the PLPL code was mainly used for self-inquiry in the form of checking ones solution.

Self – Inquiry in an Unfamiliar Context

Unlike the Feeding Time problem the Pleasant Sets problem exposed some differences in self-inquiry. Table 3 below shows self-inquiry in this context by question type.

Table 3: Self-Inquiry on the Pleasant Sets Problem by RSQ

	Low	**Middle**	**High**
Average RSQ	*-0.167*	*0.479*	*.879*
Average # of Q's	13.75	16	9.2
Definition Q's (%)	29.1	18.7	28.3
Specification Q's (%)	29.1	46.9	26.1
Legitimacy Q's (%)	41.8	34.4	45.6

The discrepancies in number of questions asked may have been expected in this context. It is possible that the High group was able to formulate the right questions to lead them to the problem solution. Meanwhile the Middle group was also able to formulate questions but had trouble formulating the correct questions to lead them to a solution efficiently. The questioning of the Low group seems to be a hybrid of the Middle and High group. The differing percentage of question types suggests further exploration of the data. Table 4 presents the data related to definition questions asked in the context of solving the Pleasant Sets problem.

Table 4: Definition Questions for Pleasant Sets by RSQ

	Pleasant Sets		
	Low	*Middle*	*High*
Avg Q's	4	3	2.6
FD (%)	0	8.3	7.7
CD (%)	81.3	75	53.8
ECD (%)	12.5	8.3	30.8
CCD (%)	6.2	8.4	7.7

The Low group asked the greatest number of definition questions followed by the Middle and High group. It is further interesting to note that the Low and Middle groups focused more than 80% of their definition questions on CD and FD questions, whereas the High group asked a higher percentage of ECD and CCD questions. It could be argued that focusing on using examples and conditions to clarify definitions is a higher order thinking skill; this was possibly the reason for fewer questions being asked by the High group as they were able to clarify their understanding of the definitions in this context with the questions asked.

The only similar discrepancies amongst Specification question types is with the Low groups focus of over 56% of their questions on CS questions as shown in Table 5 below. The most interesting difference in self-inquiry related to gathering information and understanding the problem context is the number of questions asked by the Low and Middle groups. We see that the Low and especially Middle group focus many more questions than the High group on specifics about the problem context.

Table 5: Specification Questions for Pleasant Sets by RSQ

	Pleasant Sets		
	Low	*Middle*	*High*
Avg Q's	4	7.5	2.4
CS (%)	56.2	33.3	41.6
CCS (%)	18.8	33.3	16.7
PS (%)	18.8	23.4	25
HLPS (%)	6.2	10	16.7

With regard to Legitimacy questions the self-inquiry of the three groups is similar, however we see that the High group focuses their self-inquiry on Legitimacy questions more so than Definition and Specification questions. Table 6 highlights the types of Legitimacy questions asked by the three groups.

Table 6: Legitimacy Questions for Pleasant Sets by RSQ

	Pleasant Sets		
	Low	*Middle*	*High*
Avg Q's	5.75	5.5	4.2
PL (%)	21.7	0	4.7
PLPL (%)	60.9	77.3	76.2
HLPL (%)	8.7	18.2	14.3
CFPL (%)	8.7	0	0
AFPL (%)	0	4.5	4.8

All three groups focus at least 60% of their Legitimacy questions on PLPL or "solution checking" questions. Beyond this the data seems to imply that the Middle and High group have very similar question profiles with regards to Legitimacy questions with the Middle group asking a little over a question more on average. However the Low group asked a higher percentage of PL questions implying a focus on legitimizing their problem solving procedure more so than the other two groups who focus on legitimizing their solutions.

Some Conclusions

The main goal of this project was to make a foray into understanding the self-inquiry of mathematics majors with the hope to describe self-inquiry in the context of the two problems developed. As described in the results, self-inquiry in a familiar context was relatively uniform amongst the participants, however discrepancies amongst self-inquiry seem to arise in an unfamiliar context. This unfamiliar context is similar to situations encountered as mathematics majors tackle upper-division course work. The Pleasant Sets problem requires understanding of definitions, understanding the problem context and the application of definitions for problem solution. With regards to self-inquiry on the Pleasant Sets task a description by group is plausible.

The self-inquiry of the Low group can be described as being focused at the middle tier of the question tree. The descriptions of their self-inquiry in the results suggest that the group asks a small percentage of their questions beyond the Clarification and Procedural level of the question tree. These higher-level questions seem to be what allowed the High group to efficiently contextualize definitions in this context. The main instance of questions flowing down the tree for the Low group is the use of PLPL questions, which suggest checking their problem solution. This questioning profile is highlighted in Figure 2, which depicts the question tree with words scaled appropriately by percentage of times the code was needed to classify questions.

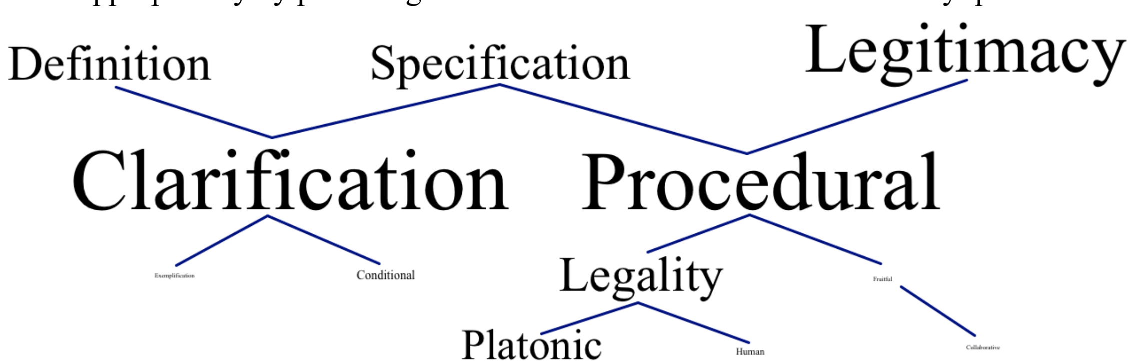


Figure 2: Self-Inquiry of the Low Group on the Pleasant Sets Problem

A similar representation of the self-inquiry of the Middle group in this context highlights the results presented that suggest the group was focused in the middle of the static to dynamic continuum described by the question tree as Specification questions dominated their self-inquiry. Unlike the Low group the Middle group focused some questions beyond the second level of tree other than PLPL questions.

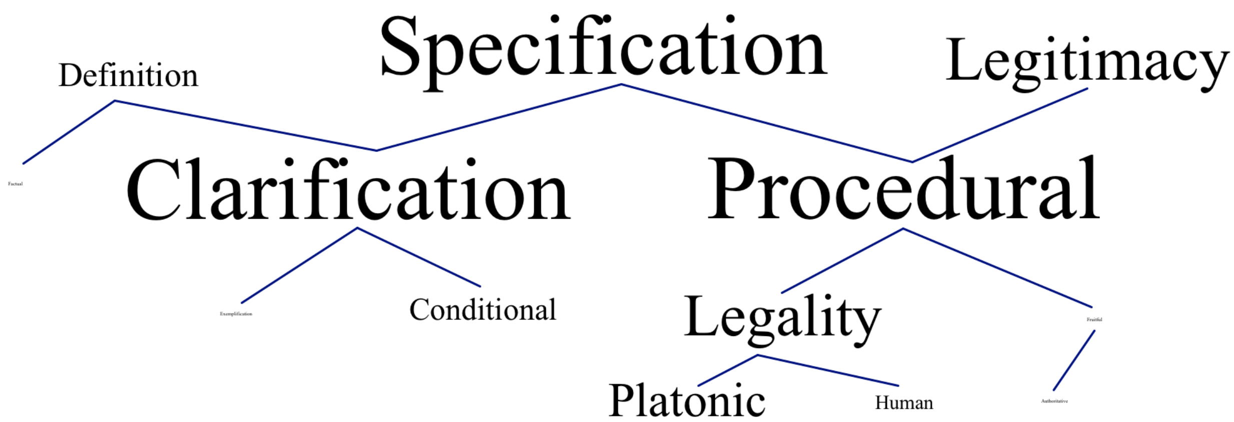


Figure 3: Self-Inquiry of the Middle Group on the Pleasant Sets Problem

The same representation for the High group implies a continued transition from the Low and Middle group to a group whose self-inquiry is focused on the dynamic side of the question tree and who utilizes questions at all levels of the tree. This was suggested in the results by the types of definition questions asked by the group and their increased focus on Legitimacy questions compared to the other two question types.

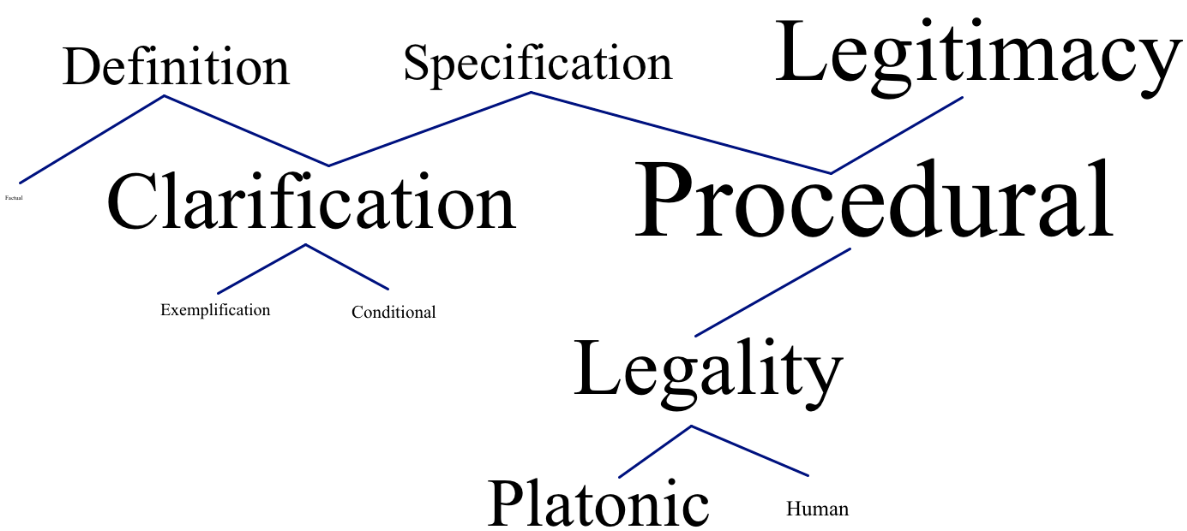


Figure 4: Self-Inquiry of the High Group on the Pleasant Sets Problem

While these descriptions are interesting they are limited to the context of the problem solving tasks presented for this data collection. However the development of the question tree and data analysis suggest potential for exploring and describing self-inquiry and hence raise a number of questions for potential future study.

Possible Future Directions

Beyond the goal of describing self-inquiry, the authors hoped to gain insight into potential next steps for a self-inquiry research agenda. First the abilities of the participants could possibly be considered a limitation of this work as the RSQ's really suggest they are all "good" students and the comparisons are made relative to the group of students. It would be beneficial to implement a comparison in which "good" could be defined by problem solving acumen on a given task. This could be accomplished by devising a set of problems for which complete solutions are not expected from all participants. It would be worth exploring if defining "good" in this manner suggests similar self-inquiry profiles.

Second the results of this work seem to suggest that the ability to understand and contextualize definitions through self-inquiry is a trademark of "good" students. This hypothesis should be tested. The above suggested work could address this but a study in which students at more varying levels of abilities were recruited may also provide valuable information about this hypothesis.

Third the data collected for this work was done in a semi-structured interview. While this seems to have been reasonably effective, other data collection techniques should be explored to determine if they make the identification of participant questions easier. For example, it is worth exploring data collection through the use of video and potentially other technologies like smart pens. These technologies may allow a more seamless linking of questions to a participant's progress being made in the problem solving process.

Finally this work has led to a number of other questions, raised by the authors, related to self-inquiry that may be worth exploring.

- Is the question tree exhaustive in the context of undergraduate mathematical problem solving?
- Would data collection with problem solving in pairs lead to better verbalization of individual thought, and therefore further conclusions about self-inquiry?
- Would data collection that utilizes new technologies provide the ability for more detailed descriptions of self-inquiry?
- Once we identify and describe self-inquiry can we develop tasks that will promote the development of self-inquiry abilities?

Potential Implications

The study described above presents data that suggests some discrepancies amongst the self-inquiry of undergraduate mathematics majors and hence provides motivation for further study of self-inquiry. The potential implications of a research agenda focused on self-inquiry are mainly pedagogical in nature. Understanding self-inquiry that leads to efficient problem solving would provide instructors with a description of question types to be utilized with students that have potential to aid student problem solving. Beyond arming instructors with new tools the main implications would be the possibility of determining ways within the mathematics major to develop students' self-inquiry abilities. Improving self-inquiry abilities may limit the number of interactions between student and teacher described in the beginning of this paper and may allow undergraduates to develop more autonomy and confidence in problem solving.

References

Belnap, J. & Parrott, A., (2013). Contributed Report: Understanding Mathematical Conjecturing. *RUME Conference, Denver CO.*, 2013.

Dillon, J.T. (1984). The Classification of Research Questions. *Review of Educational Research, 54*(3), pgs. 327-361.

Inglis, M. & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43, 358-390.

Inglis, M., Mejia-Ramos, J.P., Weber, K., & Alcock, L. (in press). On mathematicians' different standards when evaluating elementary proofs. To appear in *Topics in Cognitive Science*.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14*(4), pgs. 366-381.

Kramarski, B. & Dudai, V. (2009). Group-metacognitive support for online inquiry in mathematics with differentiated self-questioning. *Journal of Educational Computing Research, 40*(4), pgs. 377-404.

Moschitti, A., Quarteroni, S., Basili, R., & Manandhar, S. (2007). Exploiting Syntactic and Shallow Semantic Kernels for Question/Answer Classification. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pgs. 776-783.

Polya, G. (1946). *How to Solve It*. Princeton University Press, Princeton, N.J.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), Handbook for Research on Mathematics Teaching and Learning (pp. 334-370). New York: MacMillan.

Smith, N.L. & Mukherjee, P. (1994). Notes: Classifying Research Questions Addressed in Published Evaluation Studies. *Educational Evaluations and Policy Analysis, 16*(2), pgs. 223-230.

Soto-Johnson, H., & Oehrtman, M. (2012). Preliminary Research Report: Construct Analysis of Complex Variables: Hypotheses and Historical Perspectives. *RUME Conference, Portland OR.*, 2012.

VENN DIAGRAMS AS VISUAL REPRESENTATIONS OF ADDITIVE AND MULTIPLICATIVE REASONING IN COUNTING PROBLEMS

Aviva Halani
Arizona State University

This case study explored how a student could use Venn diagrams to explain his reasoning while solving counting problems. Open coding was used to identify the representations he used, and the ways of thinking in which he engaged were analyzed using an existing framework. Venn diagrams were first introduced as part of an alternate solution written by a supposed prior student. Following this introduction, the student in this study often chose to use Venn diagrams to explain his reasoning, stating that he was envisioning them. They were a powerful model for him – they helped him visualize the sets of elements he was counting and to recognize over counting. Further, he adopted the representation of the universal set in his diagrams when posing new questions and finding additive relationships between the solutions of the new and original questions. He transferred this representation to find multiplicative relationships in other problem posing situations.

Key words: Additive reasoning, Combinatorics, Multiplicative reasoning, Set Theory, Visualization

Introduction and Research Questions

Conventional wisdom often advises students to implement visualizations while they are solving novel problems. For example, Polya (1957) included "draw a picture or diagram" as one of his heuristics in *How To Solve It*. Further, Fischbein (1977) believed the coordination of conceptual schemes and intuitive representations to be essential for problem solving. The mathematics education community has demonstrated an increased interest in visualizations in mathematics, both in understanding students' visual representations and in helping these students build their intuitive visual images in order to understand abstract concepts (Alcock & Simpson, 2004; Palais, 1999; Pinto & Tall, 2002; Roh, 2008, 2009; Tall, 1991). In line with Fischbein (1977), *visual images* in mathematics are taken to be "pictorial representations of conceptual entities and operations" (p. 154). They are conceptualized images, controlled by abstract meanings. In a sense, they constitute a language – their meanings are often fairly conventionalized and they can express a wide range of ideas by using a limited method of communication. In addition, *visualization* is taken to include processes of constructing and transforming visual mental images (Presmeg, 2006). Thus, we can refer to a student's visualizations or visual images even if these representations are not physically drawn anywhere. This study examines how a student coordinated his conceptual schemes with his visual representations while solving combinatorics problems. In particular, it focuses on this student's use of Venn diagrams.

Combinatorics provides a fertile ground for investigating students' problem solving approaches. Indeed, Piaget & Inhelder (1975) contend that children's combinatorial reasoning is a fundamental mathematical idea based in additive and multiplicative reasoning and Kavousian (2008) stated that students could solve interesting and challenging problems without much prior knowledge. However, the research indicates that students of all ages often struggle with solving counting problems (Batanero, Godino, & Navarro-Pelayo, 1997; Eizenberg & Zaslavsky, 2004; English, 1993; Hadar & Hadass, 1981). Though some studies have adopted counting problems as the backdrop for research in other aspects of student learning (Eizenberg & Zaslavsky, 2004; Fischbein & Grossman, 1997; Godino, Batanero, & Roa, 2005), research on the learning and teaching of combinatorics has sparse when compared with other topics in mathematics. Shin & Steffe (2009) began to investigate how middle school students dealt counting problems based on their additive and multiplicative reasoning and determined levels of enumeration that appeared in the students' behavior: additive, multiplicative, and recursive multiplicative enumeration. Though they provided examples of students' visual representations of the elements being counting, Shin & Steffe (2009) did not focus on how students' reasoning can be expressed in their representations. Further, the problems their students encountered did not address operations more complex than permutations of five distinct elements.

Research on Venn diagrams and their use in discrete math or probability courses seems to be of two minds. On one hand, Fischbein (1977) states that Venn diagrams are powerful models that can be used to solve a wide range of problems. Indeed, some combinatorics texts (e.g. Bogart, 2000) present Venn diagrams as a visual representation for basic counting problems. On the other hand, it has been reported that many students have trouble using Venn diagrams (Bagni, 2006; Hodgson, 1996). These studies indicate that Venn diagrams may actually hinder some students' ability to solve probability problems because of their inability to visualize set expressions. This difficulty is so prevalent that that some authors have recommended the removal of the representations from basic probability courses (Pfannkuch, Seber, & Wild, 2002). In addition, some combinatorics texts (e.g. Tucker, 2002) introduce them only while solving complex counting problems such as those involving the Principle of Inclusion-Exclusion.

This case study extends the current research by investigating combinatorial reasoning in relation to a student's visual representations, particularly Venn diagrams. It coordinates his problem solving approaches with the Venn diagram representations he stated he was envisioning. In particular, this study attempts to answer the following question: How could a student use Venn diagrams to express the additive and multiplicative reasoning he employed while solving counting problems?

Theoretical Framework

The primary tenant underlying this study's design and analysis is that knowledge is not received through the senses or communication but must be actively built up by the cognizing subject (Von Glasersfeld, 1995). Under this perspective, the role of a teacher in a mathematics classroom is to orient the students' cognitive processes and assist them in their construction of their mathematical knowledge. It is imperative to define what "mathematical knowledge" then means. 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). This study focuses on the problem solving mental act. According to Harel (2008), there are two categories of mathematical knowledge: ways of understanding and ways of thinking. The reasoning applied in a particular mathematical situation – the cognitive product of mental acts – is known as a *way of understanding*. For example, the particular solution a student might provide for a task is a way of understanding as it is a cognitive product of the mental act of problem solving. On the other hand, *ways of thinking* refer to what governs one's ways of understanding and are the cognitive characteristics of mental acts. Ways of thinking are always inferred from ways of understanding. As a student progresses through a variety of tasks, certain problem solving approaches might become clear. These are ways of thinking since they are the cognitive characteristics of the problem solving mental act. This study investigates how the student's ways of thinking, or problem solving approaches in combinatorics were expressed through his Venn diagram representations.

The author developed a preliminary framework of students' robust ways of thinking about the set of elements being counted, known as the *solution set*, in basic counting problems

(Halani, 2012a, 2012b). This framework is comprised of three main categories of problem solving approaches in combinatorics. The first category, the *Odometer* category, involves the idea of holding an item, or an array of items, constant while systematically varying the other items in order to generate the whole solution set (Halani, 2012b). Students engaging in ways of thinking from the second category, *Subsets*, ultimately envision the solution set as the union of smaller subsets. The first problems solving approach in this category, *Addition*, involves thinking locally first and counting the size of a subset of the solution set before adding on the size of its supposed complement. While engaging in *Union* thinking, a student will first think globally and envision the solution set as the union of smaller subsets, which may or may not be disjoint. Students' counting methods while engaging in Union thinking vary. Some might not attend to whether the sets are disjoint and simply add the sizes of the subsets, others might employ the principle of inclusion-exclusion in some cases, and still others might then repartition the solution set before taking the sum of the sizes of these subsets. Both Addition and Union thinking are relevant to this study and are summarized in Table 1. Because their roots are in additive reasoning, they are shown in green.

Category	**Way of Thinking**	**Description**	**Example of a task whose solution could be driven by this way of thinking:**
Subsets	**Addition**	Think locally. First find the size of a subset of the solution set. Add on the size of a subset of the supposed complement of that set. Continue in this manner. Ultimately, the entire solution set should be viewed as the union of smaller subsets.	How many 3-letter "words" can be formed from the letters *a,b,c,d,e,f* if repetition of letters is allowed and *d* must be used?
	Union	Think globally and envision the solution set as the union of smaller subsets, which may or may not be disjoint. Counting methods vary – one approach when the sets are not disjoint is to then partition the set and sum the sizes of the subsets.	How many 3-letter "words" can be formed from the letters *a,b,c,d,e,f* if repetition of letters is allowed and *d* must be used?
Problem Posing	**Deletion**	Consider a related problem whose solution set contains a subset which has a bijective correspondence with the solution set of the original problem. Find an additive relationship between the original and the new solution set.	How many 3-letter "words" can be formed from the letters *a,b,c,d,e,f* if repetition of letters is allowed and *d* must be used?
	Equivalence Classes	Consider a related problem with a solution set which can be partitioned into equivalence classes of the same size – each one of which corresponds to an element of the original solution set. Find a multiplicative relationship between the solution sets.	How many permutations of *MISSISSIPPI* are there?

Table 1: Some ways of thinking about combinatorics solution sets

The last category in the framework is *Problem Posing* (Halani, 2012a) and the ways of thinking in this group involve answering a question that has not actually been asked. In

Deletion thinking, a student will consider a related problem whose solution set contains a subset which is in one-to-one correspondence with the original solution set and then find an additive relationship between the sizes of the solution sets. This approach is sometimes referred to as "Total – Bad = Good" or "Counting the Complement." In *Equivalence Classes* thinking, on the other hand, a student will consider a related problem whose solution set can be partitioned into equivalence classes, each of which is in one-to-one correspondence with an element of the original solution set, before finding a multiplicative relationship between the sizes of the solution sets. Both Deletion and Equivalence Classes are relevant to this study and are summarized in Table 1. By their very definitions, Deletion has roots in additive reasoning and is shown in green in Table 1, whereas Equivalence Classes is multiplicative and shown in red.

Methodology

Data for this study come from a teaching experiment (Steffe & Thompson, 2000) conducted at a large southwestern university in the United States. Al, a freshman enrolled in a second-semester calculus course, participated in nine teaching sessions with the researcher over a four-week period. He had no formal knowledge of combinatorics before participating in this study. Tasks for this study involved the operations of arrangements with and without repetition, permutations with and without repetition, and combinations. In addition, it is known that students do not always interpret combinatorial tasks in the same manner that the mathematical community does (Godino et al., 2005). As a result, tasks for this study were separated into two parts: a situation and a question (or questions). See Table 2 for the tasks discussed in this paper.

In line with the underlying perspective adopted in this study, much of the teaching was conducted in an effort to orient Al's cognitive processes. For example, following the completion of many of the tasks, the author implemented the Devil's Advocate instructional provocation (Halani, Davis, & Roh, this issue) by presenting alternate solutions written by supposed former students to Al for evaluation. He reinterpreted and justified the solution if he believed it to be correct, and refuted it if he disagreed. The purpose of these interventions was to create sources of potential perturbation in order to introduce a new idea, highlight strategic knowledge, or address a potential misconception. It was through these alternate solutions that many visual representations such as tables, tree diagrams, Venn diagrams, slots, and mapping diagrams were introduced.

In a pilot study, students often over counted the size of their solution sets in a large part because they had trouble visualizing the relationships between the elements they were trying to count. In many cases, when engaging in Addition or Union thinking, the students would ultimately envision the solution set as the union of subsets, but would not be aware that their subsets were not disjoint. As a result, they would often over count the number of elements in the intersection of these subsets. In order to address this over counting error, a manipulative Venn diagram activity was implemented through Devil's Advocate (Halani et al., this issue) during the fourth session of this study for task 14(vi) shown in Table 2. This intervention primarily consisted of allowing the student to work with circles cut out of translucent cellophane which could be placed on printed out Venn diagrams. The purpose of this intervention was to help the student visualize the subsets of elements being considered in both Addition and Union thinking. In this intervention, formal set theoretic language was not used. In a large part, this decision was based on the idea that students have trouble with visualizing and representing set expressions (Bagni, 2006; Hodgson, 1996). As a result, Al's natural language was adopted for use the study. For example, instead of using the term "intersection," he chose to refer to the "overlap" in the circles and the researcher used the term as well.

Session	Task	Statement
4	14(iii)	*Situation*: Suppose we have the letters *a,b,c,d,e,f* and we are forming three-letter strings of letters ("words") from these letters. *Question*: How many 3-letters "words" can be formed from these letters if repetition of letters is not allowed and the letter "*a*" or the letter "*d*" must be used, but not both?
4	14(iv)	*Situation*: Suppose we have the letters *a,b,c,d,e,f* and we are forming three-letter strings of letters ("words") from these letters. *Question*: How many 3-letters "words" can be formed from these letters if repetition of letters is not allowed and the letter "*a*" or the letter "*d*" must be used, but not both?
5	14(vi)	*Situation*: Suppose we have the letters *a,b,c,d,e,f* and we are forming three-letter strings of letters ("words") from these letters. *Question*: How many 3-letters "words" can be formed from these letters if repetition of letters is allowed and the letter "*d*" must be used?
6	16(iii)	*Situation*: A university decides that sorority names can be three-letters chosen from the following Greek letters: $\Gamma, \Delta, \Theta, \Lambda, \Pi, \Phi, \Psi, \Omega$. *Question*: How many sorority names can be formed from these letters if repetition of letters is allowed and the letter "Θ" must be used?
8	30(v)	*Situation*: Consider the word *WELLESLEY*. We will be forming "words" from these letters. *Question*: How many "words" can be formed from the letters in "WELLESLEY" if we need 4-letter words, each letter may be used multiple times, and we must use the letter "E"?
9	31(i)	*Situation*: Consider the word *MISSISSIPPI*. We will be forming "words" from these letters. *Question*: How many "words" can be formed from the letters in "MISSISSIPPI" if we need 11-letter words created by rearranging the letters provided?

Table 2: Relevant tasks implemented in the study

Venn Diagram Activity

Al was asked during the fourth session to solve the task 14(iii) from Table 2. He first argued that if "*a*" were used, it could go in three spaces, and there would be 5×4 ways to place the letters in the other slots, so that there were $5\times4\times3$ total "words" involving the letter "*a*." He then argued that there would be the same number of "words" involving the letter "*d*" for a total of $2\times5\times4\times3$. He realized that this was the same number as the total number of 3-letter "words" that could be formed where repetition is not allowed. He adjusted his solution to require that "*d*" not be allowed when "*a*" were being used, and vice versa to find a total of $2\times3\times4\times3$. When presented with task 14(iv) shown in Table 2, he engaged in Addition thinking to add on the number of "words" that allowed for both "*a*" and "*d*" to his solution for 14(iii).

He was then presented with a sheet of paper with two overlapping circles, a circle cut out of translucent purple cellophane, a circle cut out of translucent yellow cellophane, and the following alternative argument written by a supposed former student, Ian, through Devil's

Advocate. This argument was driven by Subsets thinking and involves the principle of Inclusion-Exclusion:

We will first count all of the “words” possible including the letter “*d*”, then all of the “words” including the letter “*a*”. Since “words” including both “*d*” and “*a*” would then be counted twice – once in each of those terms, we will subtract the number of “words” using both to compensate:

If the letter “*d*” is used, then the “word” can either go *d* _ _, _*d*_, _ _ *d*. For each of these, there are 5×4 ways to place the other letters since repetition is not allowed. So there are $3\times5\times4$ “words” with the letter “*d*”. Similarly, if the letter “*a*” is used, then there are $3\times5\times4$ ways to place the letters. If we sum these terms, we have $(3\times5\times4)+(3\times5\times4)$.

Now, if both “*a*” and “*d*” are used, we could have *ad*_, *da*_, _*ad*, _*da*, *a*_*d*, *d*_*a*. For each of these, there are 4 “words” we can write. So there are 6×4 “words” using both “*a*” and “*d*”. Each of these has been counted twice and we only want to count it once, so we must subtract this from out above sum: $(3\times5\times4)+(3\times5\times4)-(6\times4)=96$.

Figure 1: Ian's argument for task 14(iv) from session 4

Al’s response to this intervention was to say that he had used Venn diagrams in English, but never in math before. He was asked to reinterpret Ian’s argument in his own words and use the manipulatives to explain the solution. After some discussion, he held up the yellow cellophane and stated that there were $3\times5\times4$ “words” being represented by that circle, before waving the purple circle to say that there were the same number of “words” there. He placed both down on the paper and said that there were 6×4 “words” in the brown area which were over counted and so that amount needed to be subtracted once.

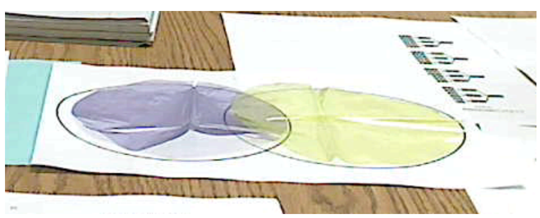


Figure 2: Two set Venn diagram activity from session 4

In the fifth session, Al was presented with another Venn diagram with manipulatives, this time with an additional circle cut out of translucent red cellophane. Since this was the second time he was encountering the manipulatives, a sheet of paper with the printed diagram was not provided. More information about this can be found in the Union section of the Results below. The concept of a universal set was not originally intended to be introduced; however, it was eventually introduced in the 8th session in order for the researcher to better understand and model Al’s problem posing ways of thinking.

There were a few phases of data analysis. Following each session with the student, content logs including a narrative of the discussions in the session and partial transcripts, along with observational, methodological, and theoretical notes (Strauss & Corbin, 1998) were created. After the data were collected, each session was transcribed. Open coding (Strauss & Corbin, 1998) was used to identify the visual representations employed by the student. In an effort understand the student’s coordination between his conceptual schemes and visual representations, his ways of thinking were analyzed using the theoretical framework from Table 1, and whether the reasoning was additive or multiplicative was also explored.

Results

After Venn diagrams were introduced during the fourth and fifth sessions of the study, Al demonstrated that he could visually represent his reasoning using Venn diagrams while engaging in the following ways of thinking: Addition, Union, Deletion and Equivalence Classes. As indicated in Table 1, the first three ways of thinking draw primarily on additive reasoning and the last draws upon multiplicative reasoning. He often employed Venn diagrams when asked to explain his reasoning and stated that he was envisioning them as he solved the tasks.

Venn Diagrams to Represent Additive Reasoning

Al recruited Venn diagrams as a tool to explain his additive reasoning employed while engaging in Union and Deletion thinking to solve counting problems.

Addition. During the fifth session, Al was asked to complete task 14(vi) from Table 2. He first over counted and found that there were $36+36+36$ "words" by arguing that "*d*" could go in one of three spaces, and for each of those options there were 6×6 ways to fill the remaining slots. First, an alternative solution driven by Deletion thinking was presented presented via Devil's Advocate (Halani et al., this issue). This intervention is discussed in Figure 7 in more depth in the Deletion subsection below. Al was asked to evaluate this argument and, after some discussion, came to find the error in his original solution. He adjusted by engaging in Addition thinking and argued that if "*d*" were first, there would be 6×6 or 36 ways to place the other letters. He then said that if "*d*" were second, there would only be 5×6 "words" he could count, and 5×5 "words" remaining if "*d*" were third. He did not visually represent this argument.

He was then presented with two more arguments via Devil's Advocate, both which relied on Venn diagrams. First, he considered an argument attributed to a former student, Adam:

> If "*d*" is first there are 6×6 ways to place the other letters. Now let's think about what happens if "*d*" is second. We already counted everything that had "*d*" first, so we can't have "*d*" first and second. Therefore, there are 5 options for the first letter and for each of them there are 6 options for the third. So there are 5×6 ways for the "*d*" to be second that we have not already counted. Finally, let's think of what can happen if "*d*" is third. We already counted everything that had "*d*" first or second, so we can't have "*d*" in either of those spots. So there are 5×5 ways to place "*d*" third that we have not already counted. Altogether we have $(6\times6)+(5\times6)+(5\times5)$ total "words".

Figure 3: Adam's argument for task 14(vi) from session 5

Al was encouraged to use the translucent cellophane manipulatives provided to reinterpret Adam's argument. This time, a sheet of paper with the overlapping circles was not provided with the intention that Al determine the alignment of the circles himself. Perhaps because his final solution to task 14(vi) was driven by the same way of thinking, Al had no trouble justifying Adam's solution with the Venn diagram. See Figure 4. Though Al did not indicate that he was envisioning Venn diagrams while he engaged in Addition thinking for task 14(vi) originally, he was able to use them to explain Adam's argument which was driven by Addition thinking.

Figure 4: Three set Venn diagram activity from session 5

Union. Al used Venn diagrams to represent his own additive reasoning for Union thinking a few times. For the task 14(vi) in Table 2, after explaining Adam's solution, Al was then asked to evaluate a third argument driven by Union thinking whose counting method involved the Principle of Inclusion-Exclusion. Again, he used the manipulatives to represent and justify this solution.

In the next session of the study during a mid-study test, Al was asked to complete task 16 (iii) from Table 2. He first drew three sets of three slots and wrote 1 in the first slot of the first set, 1 in the second slot for the second set and 1 in the third slot for the third set. Each set of slots corresponded to a different subset of the solution set, based on the location of Θ. This indicates that he was engaging in Union thinking. He filled out the remaining slots and his solution is shown in Figure 5. While determining the number of possible options in each case, Al was careful to avoid over counting by partitioning this union of subsets. He multiplied the numbers in the slots for each set and then took the sum of these products to get 64+56+49.

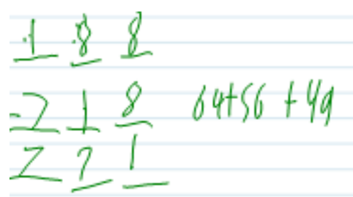


Figure 5: Al's solution to task 16(iii) in session 6

When asked about his confidence in his solution, Al referenced doing task 14(vi) during the fifth session and immediately drew a Venn diagram (not shown) to illustrate his additive reasoning. He explained his thinking:

> "*I was trying to think, ok, we have each of these different I guess groups of where it can be. Like with this one I could tell that you have a group where it's* [Θ *is*] *the first letter, a group where it's the second letter, a group where it's the third letter* (draws three overlapping circles). […] *And I knew that for all of this* (indicates all of the first set)*, I can only count this much of this* (indicates the elements in the second set excluding the first set)*, and I can only count this much of this* (indicates the elements in the third set which have not yet been counted)."

Even though Al did not draw a Venn diagram during his counting, it seems as if he may have been envisioning one from his explanation. It is clear that while he was counting, he was attending to the intersection of the subsets based on the location of Θ. The first Venn diagram Al drew was hard to read so Al drew a second one (see Figure 6) and utilized different shading techniques to show what he counted in each row of slots. In his diagram, "1st" referred to where Θ was the first letter in the "word," and so forth.

Figure 6: Venn diagram for Union thinking from task 16(iii) in session 6

When Al was asked to compare his current thinking about this type of problem to his reasoning for task 14(vi), he responded:

> "*Well, I think before, I would list them all, or I guess I didn't have as clear of a way of understanding that repetitions occur in this type of problem.* […] [*Now*] *I'm using some way to define what these three sets are. And I'm defining* […] *the first set as places where the first variable is theta. Defining that group* (points to second circle in Figure 6) *as where the second variable is theta, and that group* (points to third circle) *where the third variable is theta. And by defining them, I guess I was kind of realizing that they overlap when both the first and the second requirements are met. Or when the first and the third. Or when all three are met. So by kind of knowing that the only place I'm going to have repetitions is where that's true and that's true* (points to an intersection of two sets)*, or when all three are true, then I could kind of look for it better.*"

Here, it is clear that he was envisioning this Venn diagram even though he did not originally visually represent his reasoning while solving the task. When he refers to "repetition," he is referring to the elements of the solution set which are in more than one of his subsets, not the repetition of the letters in the words. From his comparison of his thinking while solving task 16(iii) to his thinking for task 14(vi), it appears as if Venn diagrams helped him clearly picture what he was enumerating so that he was better able to avoid over counting while engaging in Subsets thinking.

Deletion. Al employed two different variations of Venn diagrams to represent his Deletion thinking – one in the fifth session and a second during the eighth.

Session Five. As mentioned in the Addition section above, when first solving task 14(vi), Al over counted and determined a solution of 108. Devil's Advocate (Halani et al., this issue) was used to provide the following argument for task 14 (vi) – it is driven by Deletion thinking and attributed to a supposed former student, Carrie:

We first determine the number of 3-letter "words" possible regardless of whether "*d*" is used: $6\times 6\times 6$. Then, we determine the number of "words" which do not include "*d*": $5\times 5\times 5$. Thus, there are $6^3-5^3=91$ "words" which include the letter "*d*."

Figure 7: Carrie's argument for task 14(vi) from session 5

As mentioned above, Al was asked to evaluate this solution and he experienced perturbation because he felt it was correct though it yielded a different solution than his original solution. Al eventually recognized that his first answer was over counting certain elements of the solution set and adjusted by engaging in Addition thinking. After Al resolved his perturbation, he was asked if he had seen an argument like Carrie's before. The intention of the question was to address the aspects of the underlying way of thinking – Deletion – which he had naturally engaged in for several previous problems; however, his response referred to Venn diagrams:

> "*It's kind of like the Venn diagram but it's kind of not.* […] *It's kind of like the Venn diagram, cause in the Venn diagram you have kind of these two circles* (draws the two circles in Figure 8), *but she was saying that is with* 'd' (writes "*d*" in the portion in the right circle that is not in the left circle) *and then this is with all the possibilities without* 'd' (writes "~~*d*~~" in the portion in the intersection of the circles). *So she just kind of ignored this* (scribbles in the portion of the left circle that is not in the right circle)*...this is all the possibilities with* 'd,' (indicates the entirety of the right circle) *then she subtract[ed] the* [*possibilities*] *without a '*d*' to figure out how many just have* 'd'"

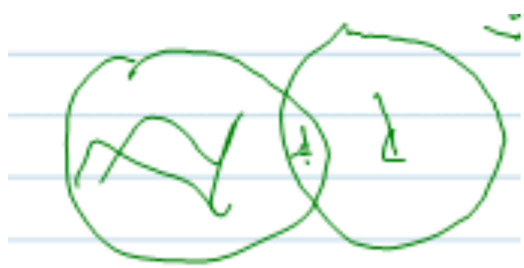


Figure 8: Venn diagram for Deletion thinking from task 14(vi) in session 5

At this point in the study, the Venn diagram activity had been implemented with two sets as seen in Figure 1: Ian's argument for task 14(iv) from session 4. In that situation, the Venn diagram involved two sets with a non-empty intersection. As mentioned in the previous section, the concept of a universal set had not been introduced and the Venn diagram activity with three sets was introduced for Adam's argument after the discussion of Carrie's argument had finished. Thus, Al's representation for Carrie's reasoning was based off the Venn diagrams he had seen before and therefore involved two sets with a non-empty intersection. This is an example of actor-oriented transfer (Lobato & Siebert, 2002).

Al's visual representation for Carrie's Deletion thinking involved counting everything in the right circle of Figure 8 and then subtracting the number of elements in the intersection. Thus, it seems as if Al understood that Carrie constructed a new problem (that of determining the total number of 3-letter words) and then found an additive relationship between the new solution set and the original solution set, even though his Venn diagram included unnecessary visual elements.

Session Eight. In the eighth session, Al tried to solve task 30(v) from Table 2. At first, he over counted and found the answer to be $\binom{4}{1}5^3$ because he considered places the E could go and then determined that there 5 choices for each of the remaining spots. The researcher reminded Al that he should ensure that he had counted everything he wanted to count and that he had not counted the same thing more than once. He quickly realized his mistake and determined the solution to be $5^3+5^2\times4+5\times16+4^3$ by engaging in Union thinking with subsets determined by the location of E and then carefully ensuring he does not over count the intersections of these subsets. He explained that it reminded him of the "*Venn diagram problem and that kind of whole picture* (draws a diagram with 4 overlapping circles) *just popped into my head*." It is not possible to draw a true diagram that shows all possible logical relations between finite sets of elements using circles and so a true "Venn diagram" for 4 sets would require ellipses or some other figures. However, to Al, this was not a factor. He was not truly envisioning all of these relations, but using the visualization to represent the fact that they exist. Once again, it is clear that he is envisioning a version of a Venn diagram for Union thinking although he did not draw it while counting.

The researcher reminded Al of Carrie's Deletion argument for task 14 (vi). Though he was not asked to do so, Al engaged in Deletion thinking for task 30(v), saying "*So in this case, it would be* 5^4-4^4." At this point, the researcher introduced the Venn diagram shown in Figure 9, explaining that the box represented the whole universe that they were concerned with. Mimicking Al's previous diagram used for task 30(v), she sketched the four circles representing subsets based on the location of E. The researcher asked Al what was actually being counted in each term of his solution. Al quickly responded that the entire box was being counted and then everything that was not in the circles was being subtracted. As before, Al demonstrated that he could use Venn diagrams to represent his additive reasoning employed while engaging in Deletion or Union thinking.

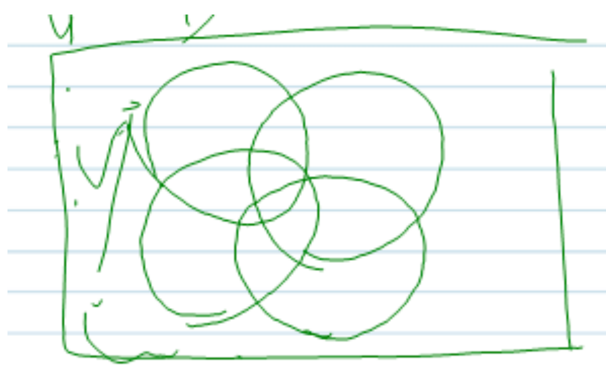


Figure 9: Venn diagram for Deletion thinking for task 30(v) in session 8

Venn Diagrams to Represent Multiplicative Reasoning

In this study, mapping diagrams and tables were introduced as visual representations for Equivalence Classes thinking, yet Al never employed them himself. Instead, Al seemed to make a connection between two Problem Posing ways of thinking, Deletion and Equivalence Classes. He used Venn diagrams with a universal set to represent his multiplicative reasoning in the latter case. This seems to be an example of actor-oriented transfer (Lobato & Siebert, 2002).

In the last session of the study, the researcher asked Al to give some examples of visual representations. His response regarding Venn diagrams is below:

> *"There's been kind of Venn diagram style overlap* (draws the Venn diagram with a rectangle and three circles shown in Figure 10) *and then there's been kind of a way that you could also figure that out by taking the whole* (indicates entire rectangle) […] *and then you're dividing out* […] *this kind of bad area* (shades in the complement of the three circles shown in grey) [...] *Because when it comes to situations with* […] *a lot of different overlaps* […] *like if there's a fourth circle* (draws the fourth circle in the figure, shown in grey) […] *then it'll get kind of complicated and so it would almost be easier to kind of find the whole thing and then kind of take out the stuff you don't want* […] [*by dividing*] […] *You figure out the ratio."*

To Al, the universal set in Figure 10 is the solution set to a different problem, one which involves both things that he wants to count, represented as the union of the circles, and things that he does not want to count. In the previous session discussed above in the Deletion section, Al determined an additive relationship between the solution set of the original problem and that of the new problem, representing the former with the universal set and the latter contained within the universal set. In his explanation about a generic problem, Al can imagine a multiplicative relationship existing instead and using the ratio to solve the problem. There are very few differences between the Venn diagrams in Figure 9 and Figure 10; however, Al was using them represent reasoning with different bases – additive in the former, and multiplicative in the latter.

Figure 10: Venn diagram for Equivalence Classes thinking from session 9

Al then demonstrated his use of Venn diagrams to represent his multiplicative reasoning in a specific problem. He had previously engaged in Equivalence Classes to solve task 31(i) from Table 2, determining that there are $\frac{11!}{4!\cdot 4!\cdot 2!}$ permutations of the letters in MISSISSIPPI. At this point, he returned to the problem and explained that there were 11! ways to permute 11 distinct items and drew a rectangle to represent these 11! elements. He then drew an oval in this rectangle, stated that we only wanted the valid answers, and wrote "*g*" for "good" inside the oval. He explained that for each "good" thing there were 4! ways to rearrange the Ss, 4! ways to rearrange the Is and 2! ways to rearrange the Ps. He stated that $4!\cdot 4!\cdot 2!$ was "*how many times more answers we have than we have valid answers*" while shading in the complement of the set "*g*." He summarized his approach:

> *"I knew if I were to attempt to try to find what's inside the 'g' by itself, it's kind of hard. But I realized that if I were able to find everything* […]*, it would be a bit easier."*

For this problem, Al first realized that he could pose a different problem – that of permuting 11 distinct objects, representing its solution set with a universal set. This concept of a universal set was something that Al seemed to connect with posing a new problem, even though it was introduced for the additive Deletion thinking in session 8, not multiplicative

Equivalence Classes thinking. He then determined that there were $4! \cdot 4! \cdot 2!$ of these elements which corresponded to each one element he actually wanted to count, representing the set of "valid answers" as a subset of the universal set. It is interesting to note that in the Venn diagram for task 30(v), the set of "words" which do not use "E" is a subset of the total number of words and were represented as such. In task 31(i), on the other hand, the set of permutations of MISSISSIPPI is *not* a subset of the set of permutations of 11 distinct items. However, there is a subset of 11 distinct items that exist in a bijective correspondence with the set of permutations of MISSISSIPPI. This subtlety did not appear to occur to Al. It is clear that he visualized a Venn diagram with a universal set containing a proper subset to explain the multiplicative reasoning he employed while engaging in Equivalence Classes for this problem.

Discussion

This study contributes to the body of research on student visualizations in mathematics. It is a step towards better understanding the connection between a student's reasoning and his visualizations in the field of combinatorics. In particular, it focuses on this student's use of Venn diagrams to represent the additive and multiplicative reasoning he employed as he engaged in various problem solving approaches.

The Venn diagram activity was introduced in sessions 4 and 5 in an attempt to address student over counting while engaging in Subsets thinking. Students engaging in both Addition and Union thinking ultimately view the solution set as the union of smaller subsets. However, their subsets are not always disjoint and, if they simply sum the sizes of the subsets, the students may over count the elements in the intersections of these subsets. Indeed, Al over counted elements of this type in sessions 4 and 5. After the implementation of the Venn diagram activity in session 5, however, he admitted that he was better able to look for the elements of the solution set which would be counted more than once. Thus, it appears as if the activity helped Al build up his mental images and sensitized him the intersection of the subsets he was considering. He did over count in session 8, but was able to recognize his mistake and adjust his solution. Again, he stated that he was visualizing a Venn diagram when he did so. Therefore, the data indicate that Venn diagrams were a powerful model for Al as he engaged in Subsets thinking – he often stated he was envisioning them as he was solving counting problems, and they helped him avoid or recognize over counting. He used them to visually represent additive reasoning in both Addition and Union thinking.

Though Venn diagrams were introduced for Subsets thinking, Al transferred the representation to the ways of thinking belonging to the Problem Posing category as well. In session 5, Al employed a Venn diagram to express the additive reasoning employed in Carrie's argument, which was driven by Deletion thinking. At that point, the Venn diagram Al had encountered involved two sets with a non-empty intersection. Thus, his representation for Carrie's thinking in Figure 8 contained visual aspects he then chose to ignore. This is an example of actor-oriented transfer (Lobato & Siebert, 2002). The concept of a universal set was introduced in session 8 in an effort to avoid these unnecessary visual aspects of his Venn diagram for Deletion thinking. Here, the universal set represented the solution set of a new problem and a subset of the universal set was the solution set of the original problem. Al then transferred the idea of the universal set to represent his multiplicative Equivalence Classes thinking in session 9. As Fischbein (1977) noted, a particular struggle for teachers and curriculum developers is to determine ways to facilitate the coordination of students' intuitive visual representations and their conceptual schema and improve their results. By understanding Al's natural inclinations in using Venn diagrams to represent his reasoning in these Problem Posing ways of thinking, we can see how this particular student coordinated such his representations with his schemes, and could use those results to design instructional activities.

The results of this case study support the inclusion of Venn diagrams in the combinatorics or basic probability curriculum as early as the use of arrangements with repetition. First, Venn diagrams may help students explain their additive and multiplicative reasoning, just as they did for Al. In addition, it seems as if introducing Venn diagrams could push students to become more cognizant of over counting while engaging in Subsets thinking and recognize how to correct these types of errors. The use of manipulatives in the Venn diagram activity could help students avoid confusion about set expressions. Further, based on how Al transferred Venn diagrams to Problem Posing thinking, it appears that it may help students to see Venn diagrams used to represent reasoning associated with these ways of thinking. The concept of the universal set could be introduced to represent the solution set of a new problem and a proper subset could be drawn inside to correspond to the elements of the original solution set. This could deepen the students' understanding of these approaches – it would emphasize the similarities Al saw between Deletion and Equivalence Classes – and could facilitate a discussion of the additive and multiplicative reasoning in these two problem solving approaches.

References

Alcock, L., & Simpson, Adrian. (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.

Bagni, Giorgio T. (2006). Some Cognitive Difficulties Related to the Represenations of Two Major Concepts of Set Theory. *Educational Studies in Mathematics, 62*(3), 259-280.

Batanero, Carmen, Godino, Juan D., & Navarro-Pelayo, Virginia. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*(2), 181-199.

Bogart, Kenneth. (2000). *Introductory Combinatorics* (3 ed.). San Diego, CA: Harcourt Academic Press.

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

English, L. (1993). Childrens' strategies for solving two- and three- dimensional combinatorial problems. *Journal for Research in Mathematics Education 24*(3), 255-273.

Fischbein, E. (1977). Image and Concept in Learning Mathematics. *Educational Studies in Mathematics, 8*(2), 153-165.

Fischbein, E., & Grossman, A. (1997). Schemata and intuitions in combinatorial reasoning. *Educational Studies in Mathematics, 34*(1), 27-47.

Godino, Juan D., Batanero, Carmen, & Roa, Rafael. (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.

Halani, Aviva. (2012a). Students' ways of thinking about combinatorics solution sets: Deletion and Equivalence Classes. . *Proceedings of the 34th Annual Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education*, 206-210.

Halani, Aviva. (2012b). Students' ways of thinking about enumerative combinatorics solution sets: The odometer category. *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education, 1*, 231-245.

Halani, Aviva, Davis, Owen, & Roh, Kyeong Hah. (this issue). Critiquing the Reasoning of Others: Devil's Advocate and Peer Interpretations as Instructional Interventions. *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education*.

Harel, Guershon. (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.

Hodgson, Ted. (1996). Students' Ability to Visualize Set Expressions: An Initial Investigation. *Educational Studies in Mathematics, 30*(4), 159-178.

Kavousian, Shabnam. (2008). *Enquiries into undergraduate students' understanding of combinatorial structures.* (Unpublished doctoral dissertation), Simon Fraser University, Vancouver, BC.

Lobato, Joanne, & Siebert, Daniel. (2002). Quantitative reasoning in a reconceived view of transfer. *Journal of Mathematical Behavior, 21*, 87-116.

Palais, Richard S. (1999). The visualization of mathematics: Towards a mathematical exploratorium. *Notices of the AMS, 46*(6), 647-658.

Pfannkuch, Maxine, Seber, George A. F., & Wild, Chris J. (2002). Probability Wih Less Pain. *Teaching Statistics, 24*(1), 24-30.

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

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.

Polya, G. (1957). *How to solve it: A new aspect of mathematical method*: Princeton University Press.

Presmeg, Norma C. (2006). Research on visualization in learning and teaching mathematics. *Handbook of research on the psychology of mathematics education*, 205-235.

Roh, Kyeong Hah. (2008). Students' images and their understanding of definitions of the limit of a sequence. *Educational Studies in Mathematics, 69*(3), 217-233. doi: 10.1007/s10649-008-9128-2

Roh, Kyeong Hah. (2009). 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*(3), 263-279. doi: 10.1007/s10649-009-9210-4

Shin, J., & Steffe, L.P. (2009). *Seventh-graders' use of additive and multiplicative reasoning for enumerative combinatorics problems.* Paper presented at the 31st Annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Atlanta, GA: Georgia State University.

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, Anselm, & Corbin, Juliet. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (Second ed.). Thousand Oaks, CA: Sage Publications.

Tall, David. (1991). Intuition and rigour: the role of visualization in the calculus (pp. 105-119): Citeseer.

Tucker, Alan. (2002). *Applied Combinatorics* (4 ed.). New York, NY: John Wiley & Sons.

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. Washington, D.C.: Falmer Press.

CRITIQUING THE REASONING OF OTHERS: DEVIL'S ADVOCATE AND PEER INTERPRETATIONS AS INSTRUCTIONAL INTERVENTIONS

Aviva Halani	Owen Davis	Kyeong Hah Roh
Arizona State University	Arizona State University	Arizona State University

This study investigated the ways in which college mathematics teachers might encourage the development of student reasoning through critiquing activities. In particular, we focused on identifying situations in which the instructional interventions were implemented to encourage the critiquing of arguments and in which students explained another's reasoning. Data for the study come from two teaching experiments – one from the domain of combinatorics and the other from real analysis. Through open coding of the data, Devil's Advocate and Peer Interpretations emerged as effective interventions for the creation of sources of perturbation for the students and for assisting in the resolution of a state of disequilibrium. These two interventions differed in design and in the type of reasoning students evaluate, but they both provoked students to further develop their reasoning, and therefore their understanding. We discuss the implications of these interventions for both research and teaching practice.

Key words: Combinatorics, Perturbation, Real Analysis, Student Reasoning, Teaching Practice

Introduction and Research Questions

The purpose of this study is to examine how mathematics teachers may leverage college students' reasoning and understanding of advanced mathematics by engaging them in critiquing the reasoning of others. Critiquing activities may perturb students who are not familiar with the ideas employed by the argument given or those whose reasoning have a different base. It is hoped that through such activities students can grasp the essence of the argument, develop their ability to distinguish correct logic from flawed, and explain the flaws if they exist. Hence activities of critiquing others' arguments would provide students a way to resolve their perturbation and to develop their mathematical reasoning. In fact, it might be difficult for a student to find flaws in his or her own reasoning. However, critiquing arguments written by others or fictional characters might enable a student to reflect on his or her own reasoning and understanding. In line with the standpoint, the recently released Common Core State Standards (CCSS, 2011) state that mathematically proficient students of all levels develop the skills to construct viable arguments and critique the reasoning of others. The mathematics education community also has demonstrated increasing interests in using critiquing activities as a research method (Lockwood, 2011; Selden & Selden, 2003) or an instructional intervention (Halani, 2012; Kasman, 2006; Roh & Lee, 2011).

This paper focuses on college instructors' creation of sources of student perturbation and their ways to facilitate students' understanding of new ideas through critiquing activities in the domains of combinatorics and real analysis. While critiquing an argument, a student must build a model of the reasoning employed in such an argument. While a body of research explored teachers' models of their students' mathematics (e.g. Courtney, 2010; Silverman & Thompson, 2008), this study emphasizes students' construction of models of their peers' or fictional characters' arguments. Focusing on teachers' role in the development of such a construction, this study addresses the following research question: How might college mathematics instructors use critiquing activities to create sources of potential student perturbations along with the ways to resolve such perturbations?

Theoretical Framework

Under the perspective of constructivism (Von Glasersfeld, 1995) adopted in this study, the role of an instructor is to orient students' cognitive processes and aid students with their construction of mathematics. One way that a teacher might do exactly this is to create sources of potential perturbation along with ways to resolve it in order to encourage students to develop their reasoning, which is what we call *instructional provocations* in this paper.

Devil's Advocate (DA) and Peer Interpretations (PI) are two instructional provocations which are designed to encourage student explanation and justification. The first, *Devil's Advocate* (DA), refers to an atypical argument provided to students by the instructor for evaluation. The idea is that instructor believes that the argument is atypical *to the student* who evaluates the argument. In fact, the argument may or may not be valid mathematically. However, regardless of its mathematical validity, the student might consider the argument to be atypical and would therefore create a source of potential perturbation. In particular, the purpose of this provocation is to highlight cognitive conflicts or to raise awareness of certain aspects of a topic. The students would refute the argument if they disagree or provide justification for parts of the argument otherwise. Instructors and researchers may design DA arguments that are attributed to a former student since the students in the class might turn a more critical eye to arguments written by their peers than they might towards arguments attributed to their instructor. The second, *Peer Interpretations* (PI), refers to a student's interpretation of a peer's argument, at the prompting of the instructor. The purpose of this provocation is to highlight similarities and differences in thinking and to allow students to learn from each other. The student interpreting the peer's argument would often include his or her own reasoning and thinking into the interpretation. PI is intended to cause perturbation in one or more of the students through the activity of highlighting of similarities and differences in thinking between students. The resulting discussion, conducted with the aid of the instructor, could help to resolve this perturbation. DA and PI hence extend Rasmussen and Marrongelle's (2006) notion of "generative alternatives" to generating student justification for the validity of alternative arguments or representations.

DA and PI are similar in that both are designed and implemented by instructors with an expectation of creating sources of perturbation or helping students resolve such perturbation. However, DA and PI differ in two aspects: First, when these provocations are designed is different – while there is often a predefined argument with using DA, PI comes about in the classroom as the interactions between students dynamically develop. Second, the types of reasoning presented are usually different. Since DA is typically prepared by the instructor prior to the lesson, DA is well-formed so that students analyze the essence of an argument. On the other hand, through PI, students analyze their peer's reasoning which is often in a formative stage so that students pull out the essence of the argument.

When analyzing a teacher's use of DA and PI in critiquing activities, we posit the following aspects of teaching are related to each other: (1) the intention of a teacher regarding student learning, (2) the tasks and materials that a teacher designs or selects, and (3) a teacher's actual teaching. The three aspects of teaching have been introduced as Didactic Triad by Thompson (2009) and adopted in this study as a model for teaching activities in which the emphasis is on students' mathematical thinking (See Figure 1). In using the Didactic Triad in this study, our focus is not on an individual part of the triad, but we rather keep in mind its relations to the other two parts of the triad. For instance, whenever we focus on analyzing what a teacher intends her students to learn, we also consider the tasks that she would give her students and the way that she would shape her teaching around those tasks so that her students would learn what she intended. Similarly, whenever analyzing the tasks or materials that the teacher designed for a critiquing activity, we also focus on how the tasks help her students learn what she intended. In addition, we also pay attention to understand how she would prepare the students for those tasks, how she manages the class around those tasks, and how she follows up on those tasks to increase the likelihood that her students learned what she intended. Finally, whenever focusing on the teacher's actual teaching, we focus on how she shapes her teaching around the selected tasks so that her students would learn what she intended. All the three aspects of teaching are situated in the teacher's understanding of her students' mathematics. We anticipate that our use of the *Didactic Triad* in this study will provide a context for addressing what her students might learn from the instruction they received from their teacher.

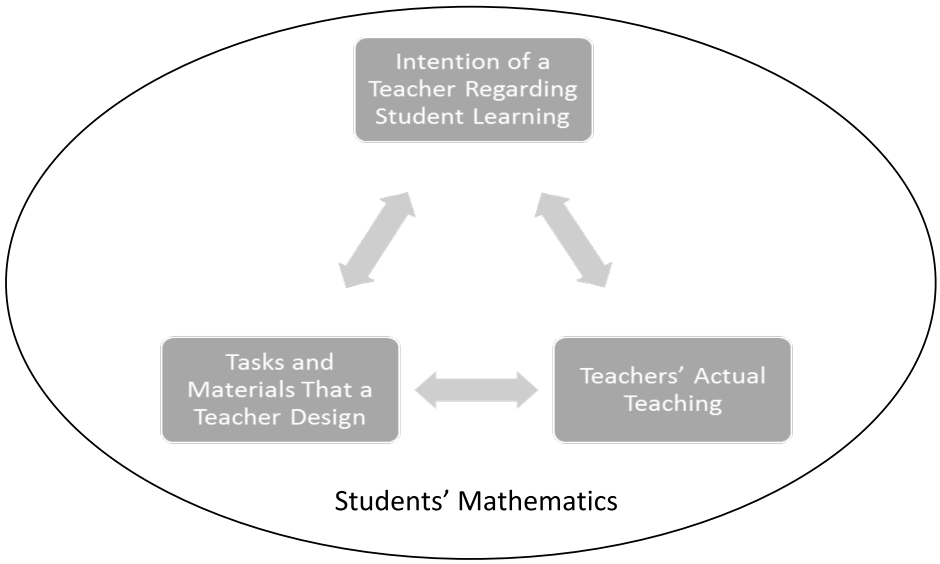


Figure 1: The Didactic Triad

Methods and Analysis

Data for this study come from two teaching experiments conducted at a large southwestern university - one from the domain of combinatorics, the other from real analysis. The rationale for including two teaching experiments is straight-forward: The content of mathematics courses typically involves the study of either discrete or continuous structures, yet in both cases, it is imperative to push students to further develop their reasoning and understanding by critiquing the reasoning of others.

Each teaching experiment constituted of 9 or 10 teaching sessions. The first author of this paper served as a teaching agent in all sessions of the combinatorics teaching experiment and some sessions of the real analysis teaching experiment, and the third author of this paper was a teaching agent in the other sessions of the real analysis teaching experiment. The second author of this paper participated in the data analysis to bring an observer's standpoint.

Both teaching experiments involved two students. In this paper we use the psuedonyms Kate and Boris for students in the combinatorics teaching experiments, and we use Jon and Sam to refer to those in the real analysis teaching experiment. The students were all high-performing undergraduate students with no prior experience in the subject. By involving two students in the teaching experiments, we wished to create a minimalist classroom where we could observe the students' interactions with each other and with the instructor.

The analysis of the data was conducted in several phases and employed grounded theory approach with open coding (Strauss & Corbin, 1998). First, we created content logs, narratives of each session, along with observational, methodological, and theoretical notes (Strauss & Corbin, 1998). Next, we transcribed all of the videos. We, as the research team, then used an open coding system to identify situations where a student explained someone else's reasoning. Following this, we used open coding to classify the instructional interventions which had the intention of encouraging students to critique the reasoning of others. The two instructional provocations, Devil's advocate (DA) and Peer Interpretations (PI) described in the theoretical framework, emerged from this phase of the data analysis. The research team then independently coded the entire transcripts and reviewed them for the consistency of the analysis.

Results

We found that in both teaching experiments, each instructor used DA and PI to push students to further develop their reasoning and understanding by critiquing the arguments of others. We found many instances for the instructors' implementation of both DA and PI throughout the study, but provide only a few illustrative examples here.

Case 1: PI in the Combinatorics Study - The ARIZONA Activity

In the combinatorics teaching experiment, the instructor, the first author of this paper, asked Kate and Boris to determine the number of ways to rearrange the letters in ARIZONA. First, Kate explained her idea:

"*I disregarded the facts that there's a repeated letter and I just said 'how many ways can* […] *you arrange these seven letters?' and that's going to be 7!. But, um, you're going to have to take some of those out.* […] *I think for every* […] *one possible order of the letters, you're going to have another* […] *that's the same because there's only one letter that is repeated. So like, if we had like just a random RZIANOA there's going to be two ways. By this, there's 7!, which count that* [RZIANOA] *twice. So I think you just divide 7! by 2 to take those out.*"

Kate determined her solution of 7!/2 by first imagining that she was permuting 7 distinct letters, though she did not use the term "distinct." She recognized that the repeated A's would actually mean that she had counted twice as many permutations as she wanted. Boris also tried to permute distinct objects first, but he tried to take away one of the A's before permuting the other 6. He then tried to insert the remaining A into the permutations he had just created, determining a solution of $6!\times 6$. The instructor asked Boris to explain Kate's argument. He responded, "*Well she went and found the total number of ways that you could arrange seven unique letters, which would be seven factorial, and she said that for each of those* […] *you're counting twice as many possibilities as you should, because of the two different A's you're assuming that those are unique letters. Like A_1 or A_2 when they're really just both A's. So you have to take out half of those*."

Notice that Boris did not repeat Kate's reasoning verbatim and instead reinterpreted it while adding further justification, thus indicating that he had built a second-order model of his peer's argument in order to extract its essence. Boris experienced disequilibrium when he realized that the two solutions the students had created could not both be correct and indicated that he believed Kate's argument to be correct by stating that he was not sure what he was counting twice. Boris eventually resolved his perturbation by recognizing mistakes in his own argument through comparing it with Kate's idea for dealing with duplicates. Thus it seems as if the instructor's request that Boris explain Kate's argument was an effective implementation of PI – it not only created a source of perturbation, but helped Boris resolve it as well.

In the next session, the students were asked to determine the number of ways to permute two blue, one red, and one black counters. Boris stated that he could think of two different ways to approach the problem. His second approach is below:

"*the other way I was thinking about doing it was the way that we, that you* [Kate] *suggested with the ARIZONA problem originally where you just do four factorial and then take out all the ones where you're double counting the blues so it would be four factorial over two.*"

Boris had assimilated Kate's way of thinking about the ARIZONA problem and applied it to this new problem. Thus, it appears as if the PI implemented during the ARIZONA problem was productive in the sense that after Boris resolved his perturbation, he was able to transfer this new way of thinking to a different problem.

Case 2: PI in the Real Analysis Study – Proofs involving Inequalities

In the real analysis study, Jon and Sam attempted to prove "For any $a,b\in \mathbf{R}$, $|a-b|\geq||a|-|b||$." In order to do so, they were directed by their instructor (the first author of the paper) towards first proving two lemmas: Let $a,b\in \mathbf{R}$, then

- Lemma 1: $|a|-|b|\leq|a-b|$ and
- Lemma 2: $-(|a|-|b|)\leq|a-b|$.

At this point, the students were already familiar with the following properties:

- Triangular Inequality (TE): Let $a,b\in$ ° , then $|a+b|\leq|a|+|b|$;
- Theorem 1 (*iii*): "Let $a,b\in \mathbf{R}$, then $|ab|=|b|\cdot|a|$"; and
- Theorem 2(*ii*): " Let $a,b,c\in$ ° , then $|a+b|\leq|a-c|+|c-b|$."

After Sam and Jon had written a proof for Lemma 1 together, they spent some time thinking about Lemma 2 separately, at the prompting of the instructor. Sam's written work for a proof of Lemma 2 can be seen in Figure 2.

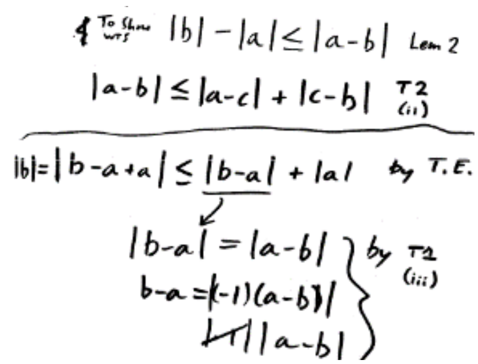


Figure 2: Sam's scratch work to prove Lemma 2: $-(|a|-|b|)\leq|a-b|$

As shown in Figure 2, Sam wrote down Lemma 2 on top of his paper, along with Theorem 2(*ii*), which he thought he might want to use in his proof. This written scratch work hinges on the idea that $|b|=|b-a+a|\leq|b-a|+|a|$ by the triangular inequality ("TE" in 2) and the use of the "adding zero technique" (not explicitly written out by Sam). The adding zero technique was a term given by the three members of the teaching experiment of adding zero to a value. For instance, one may add zero to *b* and then replace the zero with the sum of the value of *a* and its additive inverse *–a*. We see this is what Sam had before taking the absolute value to the equations such as $b=b+0=b+(-a+a)$. Sam then stated that $|b-a|=|a-b|$ was true by Theorem 1(*iii*) and by the fact that $b-a=(-1)(a-b)$. Our interpretation of Sam's proof is shown in Figure 3.

$$
\begin{aligned}
&|b|=|b+0|=|b-a+a| \text{ by "zero technique"}\\
&\Rightarrow |b|=|b-a+a|\le|b-a|+|a| \text{ by Triangular Inequality}\\
&\Rightarrow |b|\le|b-a|+|a|\\
&\Rightarrow |b|-|a|\le|b-a|\\
&\Rightarrow -|b-a|\le|a|-|b|\\
&\Rightarrow -|-1|\cdot|a-b|\le|a|-|b| \text{ by Theorem } 1(iii)\\
&\Rightarrow -(|a|-|b|)\le|a-b|
\end{aligned}
$$

Figure 3: Research team's interpretation of Sam's scratch work in Figure 2

After Sam explained his thought process, the instructor asked Jon to explain what Sam had just said. Jon indicated that he thought Sam had used Lemma 1, but Sam interrupted and pointed out that he had applied the triangular inequality instead. Because the instructor asked Jon to state his understanding of Sam's argument, Jon realized that he did not fully understand Sam's argument. In fact, in his reflection that evening, Jon wrote,

"[*The instructor*] *did ask us to explain our thinking several times, to articulate our logic. This helped me see some holes in my understanding. For example, before she asked me to explain what Sam did for Lemma* [2]*, I thought he was manipulating the inequality in order to use Lemma 1. Instead, he was using the "adding zero" technique and applying the triangular inequality in order to set up it up for Theorem* [1(*iii*)]."

This exemplifies a case when the instructor requested Jon his interpretation of Sam's argument, Jon realized that his model of Sam's argument was not what Sam intended. Since Sam indicated to Jon that Jon's model of Sam's argument was invalid, this perturbed Jon to better understand Sam's argument. The instructor's implementation of the PI was therefore effective in creating Jon's perturbation which was later resolved as the two students collaboratively worked to complete their proof of the lemmas.

Case 3: DA in the Real Analysis Study - The Vice of Inequality

At the third session of the real analysis study, the instructor, the third author of this paper, presented an alternative argument to the students in an attempt to highlight the importance of the order of quantifiers. First, she asked Sam and Jon the following question: "Would there be $x \in \mathbf{R}$ satisfying $\forall \varepsilon > 0, |x| < \varepsilon$?" After the students were given a few moments to think, the instructor asked Sam to share his thoughts. He responded,

"*Okay, so I was thinking that the only x that will work for this* […] *would be 0, because* […] *you could get x really small* (pinches fingers together)*, give it a really small value, but it's still not gonna work for any* ε *greater than 0 because the limit of that is 0. So, it's* [*x* is] *always going to be infinitessimally larger than 0, which means it* [ε] *can always be smaller than any positive x* [which is a contradiction]. *So, it would only work for 0.*"

The instructor then presented an alternative solution to the given statement, asserting that there are infinitely many possible values of *x* based on the following theorem: If *x* is between $-\varepsilon$ and $+\varepsilon$, then $|x| < \varepsilon$. The alternative argument ε results in infinitely many *x* values, while Sam responded that 0 is the only solution that satisfies the given statement. The instructor asked Sam to explain his reasoning about the alternative argument and he responded that the alternative argument cannot be a valid argument for the given statement. When pressed to discuss his reasoning for how he could tell, Sam had difficulty in doing so. This difficulty caused the perturbation necessary for Sam to create models of both the given statement and the instructor's alternative argument. After prompting from the instructor Sam presented his model of the given statement as there is "*one value of x, for which the value of |x| is always less than* ε." His representation of the instructor's alternative argument was that "*basically, you pick some value of* ε *and then it tells you for what values of* x, (that) $|x| < \varepsilon$." Once he built these models, Sam resolved his perturbation. When prompted by the instructor, Sam was able to describe the difference between the two statements, in which the order of the quantifiers had an impact on the meaning of the statements. The instructor's introduction of the alternate argument raised a cognitive conflict that was resolved by the student Sam, which indicates the instructor's use of the alternative argument was an effective use of DA.

Case 4: DA in the Combinatorics Experiment - Tree Diagrams

In the combinatorics teaching experiment, the instructor also often provided alternative arguments to the students for evaluation. One example where she did so was when Kate and Boris were asked to solve the following task, which is adapted from Batanero et al.'s (1997) questionnaire:

Situation: Four children: Alice, Bert, Carol, and Diana go to spend the night at their grandmother's home. She has two different rooms available (one on the ground floor and another upstairs) in which she could place all or some of the children to sleep.

Question: In how many different ways can the grandmother place the children in the two different rooms?"

Boris determined the answer to be 2^4, explaining that there were two rooms that the first person could go to, for each of those possibilities, there were two possibilities for where the second person could go, and so forth. Then the instructor provided the tree diagram shown in **Error! Reference source not found.** Figure 4 as a solution provided by a supposed former student, Annette.

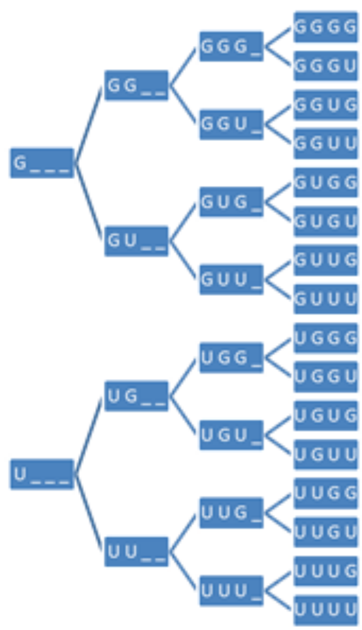


Figure 4: "Annette's Solution" provided through Devil's Advocate

At first Kate was confused by the representation and stated, "*I don't even know what that means*." After examining the tree diagram for a while, Boris stated,

"*So I guess it's like doing it per person.* […] *She [is] pulling it apart like one person at a time. For the first person, they can either go to the ground floor or the upper floor. So like, you hold one constant. Say the first goes to the ground floor.* […] *And then the next person could go to the ground floor or the upper floor. So then, they both go to the ground floor for those* […] *four possibilities* (points to the top four leaves of the tree). *After that point* (points to the vertex G G _ _) *they* [*the third person*] *can go to the ground floor or the upper floor. So if they go to the ground floor* […] *and again there are two more possibilities for each of those. So there's two more there*."

Boris had made a connection between Annette's solution and the idea of holding something constant. He was able to pull out the essence of Annette's argument and explain it in his own words. Following Boris' interpretation of Annette's solution, Kate immediately responded, "*so this is just a graphic representation of what you* [Boris] *were saying*." This indicates that Kate, as well, was able to grasp the essence of Annette's solution and connect it to Boris' original solution even though she originally experienced some perturbation and did not immediately understand the tree diagram. The instructor's intention in providing Annette's solution was to raise awareness of the existence of visual representations for their current ways of thinking. Since the students were successful in building connections between Annette's solution and Boris' original solution, therefore further developing their reasoning and understanding, we consider the instructor's introduction of Annette's solution to be an effective implementation of DA.

Later in that session, Kate and Boris were attempting to determine the number of 3-letter "words" that could be formed from the letters *a*, *b*, *c*, *d*, *e*, and *f* if repetition of letters were allowed and the letter "*d*" must be used. They first over counted and found the answer to be $3\times6\times6$. The instructor provided a DA that determined the solution to be $6^3 - 5^3 = 91$. The students realized that both solutions could not be correct but they both had trouble identifying which solution was correct and which involved a flaw in reasoning. Boris and Kate used the tree-diagram in Figure 5to solve the problem using a third method and confirm that the alternative solution provided was correct.

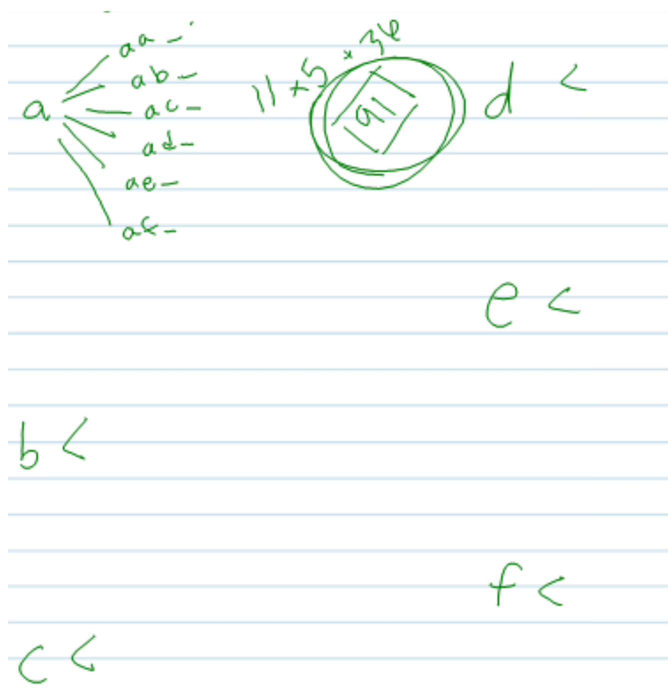


Figure 5: Kate and Boris's Transfer of Tree Diagrams

Their tree-diagram differs vastly from the one supposedly written by Annette in the earlier task – the leaves in Figure 4 each represent an element of the solution set, but in Figure 5all of the leaves are missing, many of the trees have only a root, and the use of slots to indicate where other items would be placed is inconsistent. However, Annette's idea of using a tree diagram to visually represent the elements being counted was adopted by the students. This seems to be an example of actor-oriented transfer (Lobato & Siebert, 2002).

Discussion

In this study, we found that both Peer Interpretations and Devil's Advocate were effective instructional interventions designed to encourage students to critique the reasoning of others. Such activities challenged the students to understand the mathematics of a peer, fictional or otherwise, and provided opportunities to deepen conceptual understanding. Indeed, these provocations could create sources of perturbation or assist in the resolution of such perturbation. In some cases, like in PI example from the real analysis study, the provocation may simply accomplish one of these tasks. In other cases, like both combinatorics examples and the DA example from real analysis, the provocation could both create the perturbation and help with its resolution.

Four classroom situations where DA and PI could be applied were identified over the course of this study. The first, identified as Type 1 (Figure 6), involved a valid argument that was presented in one of two ways, dependent on whether DA or PI was being used. In DA, the instructor presented the argument as a former student's argument, and one (or both) of the students interpreted the argument. In PI, one student in the class presented the argument, and the instructor asked the other student to interpret this argument. It is important to note that the validity (or invalidity) of an argument is not usually apparent to the students participating in these situations, but the instructor should be, and needed to be, aware of the validity of the arguments. In the second situation, identified as Type 2 (Figure 6), an invalid argument was presented. This invalid argument was presented through DA as a former student's argument. On the other hand, this situation did not come up with PI. The instructors in the teaching experiments indicated they did not want to hurt the feelings of the student presenting the invalid argument by having another student interpreting and critiquing the invalid argument. The Type 3 (Figure 6) situation involved two valid arguments that were presented in the classroom. With DA, the instructor presented a "former student's" valid argument, while one (or both) of the students presented their own valid argument. The students then interpreted the DA argument. In the PI situation, both students presented valid arguments and interpreted the other student's argument at the prompting of the instructor. In this situation, the students would compare the argument that was not their own to their own argument to determine the similarities and differences between the two arguments. Finally, in the Type 4 (Figure 6) situation, one argument was invalid while the other was valid. If a student presented an invalid argument, an instructor would often present a valid argument through DA, to give the student(s) an opportunity to develop models of their own argument and the DA argument. After developing these models, the students could compare and contrast these two arguments to determine which of the arguments are useful in solving the problem at hand. This is also true with the reverse situation of an invalid DA argument and a valid student argument. The students would interpret the invalid argument and compare it to the valid argument presented by the student in class. By comparing these two arguments, the students could become aware of issues brought up in the invalid DA argument that they may not have been aware of otherwise. In Type 4 PI, one student would present a valid argument and the other student would present an invalid argument. The instructor would prompt the two students to interpret each other's argument, which again has the purpose of causing students to become aware of the features of a valid argument that satisfies the given problem. Other combinations of arguments (invalid/invalid, for example) are possible, but these four types were the types identified over the course of this study.

Figure 6: Four implementations of DA & PI uncovered in this study

This study has implications for both research methods and teaching practice. As shown, both DA and PI were effectively implemented in teaching experiments (Steffe & Thompson, 2000) and provided opportunities for further discourse, thus allowing the researchers to better understand the students' reasoning. We contend that DA could be implemented in clinical interviews (Clement, 2000) in a similar manner to help the researcher confirm his or her model of the student's mathematics. Because it is possible that a student's mathematics may change as a result of DA, we recommend the use of such provocation at the end of an interview. In a classroom, a teacher could implement either DA or PI to highlight differences in reasoning and raise or resolve cognitive conflict. In both cases, the reinterpretation by a student can include the student's own thinking and reasoning, while also including his or her own interpretation of the original argument. The interpreting student may adopt the meaningful aspects of the other argument into their own model of the situation. Indeed, we found evidence of this adoption in the student's assimilation of the idea to new situations in both combinatorics episodes discussed in this paper. Both DA and PI were effective in pushing students to further develop their reasoning by critiquing the reasoning of others.

References

Batanero, Carmen, Godino, Juan D., & Navarro-Pelayo, Virginia. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*(2), 181-199.

Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In R. Lesh & A. Kelly (Eds.), *Handbook of research methodologies for science and mathematics education* (pp. 341-385). Hillsdale, NJ: Lawrence Erlbaum.

Courtney, Scott. (2010). *Exploring teachers' capacty to reflect on their practice: an investigation into the development of mathematical knowledge for teaching.* (Unpublished Doctoral Dissertation), Arizona State University.

Halani, Aviva. (2012). Students' ways of thinking about combinatorics solution sets: Deletion and Equivalence Classes. . *Proceedings of the 34th Annual Conference of the North*

American Chapter of the International Group for the Psychology of Mathematics Education, 206-210.

Kasman, Reva. (2006). Critique That! Analytical Writing Assignments in Advanced Mathematics Courses. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 16*(1), 1-15.

Lobato, Joanne, & Siebert, Daniel. (2002). Quantitative reasoning in a reconceived view of transfer. *Journal of Mathematical Behavior, 21*, 87-116.

Lockwood, Elise. (2011). *Student approaches to combinatorial enumeration: The role of set-oriented thinking.* (Unpublished Doctoral Dissertation), Portland State University, Portland, OR.

Rasmussen, Chris, & Marrongelle, Karen. (2006). Pedagogical Content Tools: Integrating Student Reasoning and Mathematics in Instruction. *Journal for Research in Mathematics Education, 37*(5), 388-420.

Roh, Kyeong Hah, & Lee, Y. (2011). The Mayan activity: A way of teaching multiple quantification in logical contexts. . *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 21*(8), 1-14.

Selden, Annie, & Selden, John. (2003). Validations of proofs considered as texts: Can undergraduate tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*, 4-36.

Silverman, J., & Thompson, P.W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education, 11*, 499-511.

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, Anselm, & Corbin, Juliet. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (Second ed.). Thousand Oaks, CA: Sage Publications.

Thompson, P.W. (2009). Case Modules Facilitator Guide.

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. Washington, D.C.: Falmer Press.

THE ACTION, PROCESS, OBJECT, AND SCHEMA THEORY FOR SAMPLING/SAMPLIING DISTRIBUITON[1]

Neil J. Hatfield
Arizona State University

This paper offers a theoretical perspective for students' understanding of sampling, samples, and sampling distributions by melding aspects of the Action, Process, Object, and Schema theory and Saldanha's and Thompson's Multiplicative Conception of Sampling. This theoretical perspective provides one potential way to describe the development of a student's conception of sampling. Additionally this perspective differs from most other perspectives in that it does not focus on the sample size the student uses or the sampling method, but rather how the student understands sampling in terms of a sampling distribution.

Key words: sampling, sampling distribution, multiplicative conception of sampling, APOS Theory, theoretical perspective

Introduction

Statistics is one of the most ubiquitous branches of mathematics in everyday life and it is also arguably one of the least understood areas of mathematics. One important aspect of Statistics is that of statistical inference. While students are taught how to calculate a statistic and then perform an appropriate test on that statistic, often the meanings that these students have is procedural (Lipson, 2003). Students rarely understand that when making inferences from a statistic, they are making judgments from a sampling distribution (theoretical or experimental). Sampling and sampling distributions are thus an integral part of statistical inference. The National Council of Teachers of Mathematics (NCTM, 2000) noted that students should be able to construct and use sampling distributions in regards to statistical inference. The Common Core State Standards for Mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) expects students to be able to make inferences about a population using a random sample and use sample statistics to make decisions. Each of these expectations requires that students develop productive ways of thinking about samples, sampling, and sampling distributions.

While there are several articles on students' understandings of sampling, the work of Saldanha and Thompson (2002) was highly influential in the development of this theoretical perspective of a student's understanding of sampling/sampling distribution. Saldanha and Thompson's multiplicative conception of sampling (MCS) serves as a basis for building a more nuanced framework for student thinking about sampling. APOS theory's applicability to a wide range of mathematical concepts including, but not limited to functions, infinity, limits, mean, standard deviation, and the Central Limit Theorem, suggests that it could provide a beneficial framework to refine MCS. In essence, this proposed theoretical perspective attempts to use elements of APOS theory to extend Saldanha and Thompson's MCS. Further, this proposed framework serves as a starting point in developing a hypothesis for how an individual might

[1] Research reported in this article was supported by National Science Foundation Grant No. DUE-1050595. Any recommendations or conclusions stated here are the authors and do not necessarily reflect official positions of the NSF.

develop particular conceptions of sampling, sample, and sampling distribution. Finally, the proposed framework might lead to further research to develop 1) descriptions of the meanings, understandings, and ways of thinking about sampling/sampling distribution that individuals may have, 2) hypotheses on the development of those meanings, understandings, and ways of thinking, and 3) how we might support students in their development of productive ways of thinking about sampling/sampling distribution.

To layout this APOS Framework for Sampling/Sampling Distribution (APOS-S/SD), a brief description of the constructs involved in APOS-S/SD, APOS theory in general, and the multiplicative conception of sampling will be presented. Following this will be a presentation of the APOS-S/SD theoretical perspective with examples.

Constructs and Background Theory

Two important terms to define for this perspective are construct and catalytic construct. Wilson (2005) described a *construct* as "part of a theoretical model of a person's cognition" (p. 6). More explicitly, a construct is an idea existing in the mind of an observer about how another individual might understand a particular concept. The observer infers the other individual's understanding of a concept through the actions/behaviors of the individual. The characterization/description of the inferred understanding constitutes the construct. Thus, the construct is part of the observer's thinking about a latent variable that resides in the other individual. Constructs refer to understandings (in the moment) or stable meanings that an individual is hypothesized as having. A *catalytic construct* serves as an observer's model of how another individual's understanding or meanings might advance within a given framework. An observer's model of the mental processes that an individual might enact would then constitute a catalytic construct.

Embedded within APOS theory, three catalytic constructs play an important role; Sfard's (1992) constructs of interiorization, condensation, and reification serve as catalytic constructs. While each of these catalytic constructs have been used in other instances of APOS theory, the meaning and intent behind each is sometimes unclear. As such, I offer an alternative meaning for each of these constructs that I use in the development of APOS-S/SD. *Interiorization* refers to an individual's mental "process performed on already familiar objects" (Sfard, 1992, p. 64). The alternative meaning I propose is very much in line with Sfard's description. Here the observer hypothesizes that an individual carries out a sequence of steps on an entity (such as a number) that the individual views as something he/she may manipulate. For Sfard (1992), *condensation* refers to taking a mental process and "turning this process into a more compact, self-contained whole" (p. 64). An issue with this characterization of condensation is that it does not offer a hypothesis for *how* this proposed transformation occurs. The alternative meaning is to note that the observer hypothesizes that the individual envisions a sequence of steps and the individual draws upon his/her experiences working through that sequence to anticipate the result of the sequence without explicitly carrying out the sequence. The individual may achieve this through a cycle of repeatedly carrying out the steps of the sequence and packing/unpacking the steps until the individual begins to anticipate the end product of any step and of the whole sequence. The final catalytic construct, *reification,* is perhaps the most problematic of the three. Reification refers to an observer's model for how an individual might "transition from an operational to a structural mode of thinking" (Sfard, 1994, p. 54). Dubinsky, Weller, McDonald, and Brown (2005a) use the term *encapsulate* in place of reification; however, their usage of encapsulate is consistent with Sfard's usage of reification. For either term, reification or encapsulate, little light is shed on what mental processes might be unfolding in the mind of the individual. An alternative would be where the observer hypothesizes that an individual has reflected upon 1) his/her experiences with the sequence of steps, 2) the anticipation of the end results, and 3) unpacked/repacked the sequence of steps so that the reflection leads the individual to ascribing characteristics that allow for manipulation to the sequence rather than the end results of the sequence. The notion of reflection is critical to this framing; the individual is reflecting upon at least three related sets of experiences. Notice that it is not enough for the observer to merely hypothesize that the individual has (or is) reflecting on the enumerated aspects; the individual must ascribe characteristics to the sequence. A distinction must be made between *characteristics* and *characteristics that allow for manipulation*. An individual may look at the rule for a function and remark, "That looks really nasty to deal with." While that individual has given a characteristic to the function ("nasty looking"), this characterization is not manipulatable in the sense that the individual sees the function as having properties that may be altered through the use of operations (from the observer's perspective). To better understand these catalytic constructs, examples will be given in the following section in the context of APOS theory and functions.

APOS Theory

Multiple researchers have used APOS theory in an attempt to describe how an individual might understand a collection of different mathematical topics (for example, see: Breidenbach, Dubinsky, Hawks, & Nichols, 1992; Clark, Kraut, Mathews, & Wimbish, 2003; Dubinsky & McDonald, 2002; Dubinsky, Weller, McDonald, & Brown, 2005a, 2005b; Mathews & Clark, 2007; Sfard, 1992). To help explain the three catalytic constructs and provide additional background for the current theoretical perspective, we will use APOS-Function[2] as a backdrop. In order for an observer to make any claims about a mental process, the observer must first specify what serves as the object(s) that the individual views as familiar. In the case of functions, we will specify (and assume) that numbers are familiar objects for the individual. Now it may be that the individual is able to manipulate numbers through the use of arithmetic operations. These arithmetic operations serve as proxies for the individual mental steps that the observer hypothesizes the individual working through. The individual manipulates the numbers according to some sequence. To the observer, the hypothesized sequence is the catalyst that enables a student to move from precursors of a meaning for function to a meaning for function that is consistent with the construct of action. *Action* refers to a conception of a concept where an individual's meanings lead him/her to following an explicit sequence of steps to transform an object. In the case of function, the action conception refers to how an individual might see a function rule as a recipe to follow in changing a number into a different number.

As an individual continues to work with a sequence of steps, the individual might begin to anticipate the result of particulars steps. For example, an individual might begin to anticipate that the step $2*x$ as doubling the value of x. As the individual begins to anticipate the result of more steps, he/she may begin to anticipate the result of a sequence of steps. Repeatedly working with steps and anticipating the results of the steps is part of condensation. When an individual begins to see functions as "self-evaluating expressions" (Thompson, 1994, p. 6), we can characterize that individual's understanding of function as being the process conception. The *Process* conception involves the individual thinking about performing the same sequence of steps without having to carry out the steps explicitly and anticipating the end result (Dubinsky & McDonald,

[2] The author uses APOS-(Concept) as a shorthand reference to the body of work where APOS theory has been applied to a specified concept.

2002). In the case of function, an individual with the process conception of function can anticipate that $f(2)$ is the end result of applying the sequence of steps denoted by f on the value 2. Further, an individual with the process conception is able to note that $f(x)$ is an end result much like $f(2)$.

The distinction between an action conception of function and a process conception of function partially involves a difference in vantage points. For the action conception of function, the individual appears to have a vantage point of *inside* the sequence of steps to the observer. In the process conception of function, the individual has a vantage point of *outside* the sequence of steps to the observer. Consider the following: $f(x) = 2x^2 - 4$. With a vantage point of inside the sequence, $f(2)$ is seen as the following sequence: square 2, get 4; multiply 4 by 2, get 8, subtract 4 from 8, get 4; the result is 4. When one is at the outer vantage point, the individual sees the whole process as one thing, i.e. $f(2)$ is what one gets out of f when one puts 2 into f.

As an individual continues to work with different sequences of steps, anticipating the end results of those sequences, the individual may begin to notice characteristics about the sequences. These sequence characteristics could serve as pointers to the individual that the sequence of steps is itself an object much like the number he/she applies the sequence of steps to or the value he/she gets (anticipates getting) after applying the sequence. The noticing of these characteristics may result from the individual reflecting upon experiences with the sequence(s) and unpacks/repacks his/her understandings of a particular sequence. Here the reflection, unpacking, and repacking serve as parts of the mental process the observer hypothesizes about the individual's cognition with regard to a particular concept. In terms of function, when the individual's meanings enable him/her to think about a sequence of steps in such a way that he/she describes manipulatable characteristics, we may say that that individual has an *object* conception of function. It is at this point that the individual may now begin to manipulate a function, say f, through the alteration of the characteristics he/she noticed about f.

Just as part of the distinction between action and process conceptions of function involves different vantage points, there is also a slightly different vantage point for the object conception of function as compared to the process conception (or even the action conception of function). With the object conception, the vantage point of the individual is also outside of the sequence of steps to the observer, but this vantage point is outside of the sequence at a much larger distance. An analogy for these vantage points lies within the notion of the function machine. The inside vantage point would place the individual's thinking within the machine; watching each step alter the input. The outside vantage point is just outside of the function machine. The individual sees the input, the function machine, and sees the output. At any point, the individual may open up the function machine and peer inside, but the individual does not need to always be looking inside of the function machine. Finally, the distant outside vantage point could be from the rafters of the building looking down at the function machine. Here, the individual cannot necessarily see the input and output of the function machine, but rather sees the function machine as something that may be moved around (or out) of the building. Even though, the individual's vantage point may be hypothesized at this distant point, the observer hypothesizes that the individual can still change to the closer outside vantage point or even return to the interior vantage point.

The final element of APOS theory refers to the individual's schema. An individual's *schema* is that "individual's collection of actions, processes, objects, and other schemas which are linked by some general principles to form a framework" in that individual's mind (Dubinsky & McDonald, 2002, p. 3). This is very similar to Tall and Vinner's construct *concept image*; a complete cognitive structure an individual associates with a particular concept (1981). While there is a hierarchy to the action, process, and object conceptions, the schema element of APOS theory is not in that hierarchy. At any point in time, an observer can hypothesize the individual's collection of cognitive elements (the understandings, meanings, ways of thinking in the sense of Thompson (2013)). Given this view of schema, by hypothesizing that an individual has an action conception of some concept is to suggest what that individual's schema might consist of. Thus, moving forward from this point, any hypothesis about an individual's conception is understood to be reflective of the description of that individual's hypothesized schema.

Multiplicative Conception of Sampling

The foundation for the APOS-S/SD framework is the multiplicative conception of sampling (MCS) as framed by Saldanha and Thompson (2002). Drawing from the work of Inhelder and Piaget (see Saldanha and Thompson's article for references), the authors ground MCS in the notion of multiplicative reasoning, which requires viewing multiple quantities/attributes of the same object simultaneously. To further explicate MCS, the authors initially stated that "having conceived a sample as a quasi-proportional mini version of the sampled population, where the 'quasi-proportionality' image emerges in anticipating a bounded variety of outcomes, were one to repeat the sampling process" (Saldanha & Thompson, 2002, p. 266) is a sampling conception that they will refer to as MCS. The multiplicative conception of sampling then consists of three levels in which multiplicative reasoning plays a key role. Saldanha and Thompson (2002) do not propose these levels in a developmental sense. In fact, it was not the intention of Saldanha and Thompson to suggest a developmental trajectory for the multiplicative conception of sampling. Rather their intent was to merely describe different conceptions of sampling. However there is a certain, natural ordering to their levels. The first level (L1) an individual being able to visualize "a relationship of proportionality between a sample and a population" (Saldanha & Thompson, 2002, pp. 266–267). The second level (L2) requires that an individual begin to see a sampling distribution begin to take shape as he/she examines a common statistic across many samples. The final level (L3), an implied ultimate level in this framework, is a view that "supports quantifying the expectation of a particular kind of sampling outcome and thus quantifying one's confidence in a sampling outcome's representativeness" (Saldanha & Thompson, 2002, p. 267). It should be noted that the observer seeking to characterize an individual's conception of sampling with these levels must be able to describe the individual's conception of sample and population. These two ideas play key roles in the three levels.

To help provide some clarification of the levels, it is helpful to look at an example. Say that there is a population of albatrosses composed of 30% Royal albatrosses and the remaining 70% are black-footed albatrosses. At L1, an individual recognizes that if he or she were to take a sample of size n, that the sample would have some relationship between its proportionality of albatross breeds and the population's proportionality of albatross breeds. There are better and worse conceptions of the relationship here; for example, a worse conception would be where the individual believes that the proportion of albatrosses in the sample is exactly the same as the proportion of albatrosses in the population. A better conception would be one where the individual acknowledges that the proportion of albatrosses in the sample is connected to the proportion of albatrosses in the population but the two proportions may differ in value. In L2, the individual begins to construct a sampling distribution. In the current example, the individual has conducted multiple samples of size n and begins to compare values of the same statistic, in this case the proportion of Royal albatrosses. As he or she compares those values, he/she clusters those values and forms hypotheses about the collection of values. While the individual actually constructs a sampling distribution of the values of the statistic, it may not be a carefully constructed, formal distribution of the values of the statistic. Finally, in L3, the individual has reached a point where the sampling distribution constructed in L2 crystalizes into a 'formal' distribution that he/she can now use to judge the representativeness of the proportion of Royal albatrosses from a new sample of size *n*. Furthermore, he/she is able to also speak about the confidence that he/she makes the representativeness judgment with.

Also within their work, Saldanha and Thompson (2002) explicate three phases that represent the sampling process. First, an individual randomly selects *n* objects from a population and calculates/records a desired statistic. Then, the individual repeats this process a desired number of times (a large number of times); each time calculating and recording the value of the same statistic. The final phase is then partitioning the collection of values for the statistic to make a desired judgment. While their work was dealing with proportions of dichotomous populations, the phases of Saldanha and Thompson's sampling process are general enough for sampling beyond their scope.

Action, Process, Object Conceptions of Sampling/Sampling Distribution

As stated previously, the intent of this framework is to propose a hypothesis for what a more developmental view of to Saldanha and Thompson's MCS might look like. APOS theory provides some opportunity to theorize what developmental details could look like. The theorizing of additional details within APOS-S/SD framework does not necessarily imply that APOS-S/SD is a developmental framework in-and-of itself. Such a framework would need to be guided by a microgenetic analysis of learning such as that suggested by Simon et al (2010). Rather, this framework could serve as a potential stepping-stone towards such an endeavor.

The APOS-S/SD framework begins with several key assumptions that must be made before attempting to describe an epistemic individual's conception of sampling and sampling distribution. First, the individual must have a conception of population and sample. For the purpose of APOS-S/SD, we will assume that the individual conceives of a population as a collection of all possible objects/people that he/she might wish to know something about. A sample then will be a randomly drawn subset of the population of a particular size with an assumption that the sample will be potentially representative of the population. Additionally, the individual needs to have an understanding that stochastic processes are repeatable processes with the anticipation that the results of the processes will vary within some bounds each time the process is repeated. Further, it must be assumed that the individual has some conception of statistic such as seeing a statistic as way to produce a measure of some attribute of the collection of objects/people. For now, we will note that all samples are of the same size. With these assumptions in place, the action conception of sampling/sampling distribution may be described.

Action Conception of Sampling/Sampling Distribution

Starting from the prior assumptions, the objects/people (elements) that constitute the population serve as the familiar objects on which the individual will act. The individual might conceive of the following sequence of steps: envision/carry out a stochastic process of drawing a sample from the population of interest, conceive of a statistic of interest, find the value of that statistic based upon the elements of the sample. For the individual to see the statistic of interest as valuable necessitates that the individual have an L1 conception in the multiplicative conception of sampling; there is a relationship between the value of the statistic stemming from the sample and the parameter in the population. The actual implementation of the drawing the sample maybe done in real life or may be part of a simulation. The student does not actually need to physically draw the sample. This would be comparable to a student using a calculator to apply the sequence of steps in a function rule. If the observer hypothesizes that the individual's understanding is best described by the individual repeatedly moving through each and every step of the sequence and the individual requiring data values to arrive at a value of the statistic, then the observer could refer to this conception of sampling as the *action conception of sampling/sampling distribution*.

Saldanha and Thompson (2002) provided several excerpts to highlight different conceptions of sampling/sample distribution. Within the context of their teaching experiment, Saldanha and Thompson (2002) had the students engage in an activity with a simulation where students worked with a computer simulation. The simulation was designed to randomly draw samples from a population of people and report back whether or not each person liked or disliked Garth Brooks's music. The interface (see Figure 1) shows that the proportion of individuals who liked Garth Brooks's music in the population is 30% (this is the population parameter).

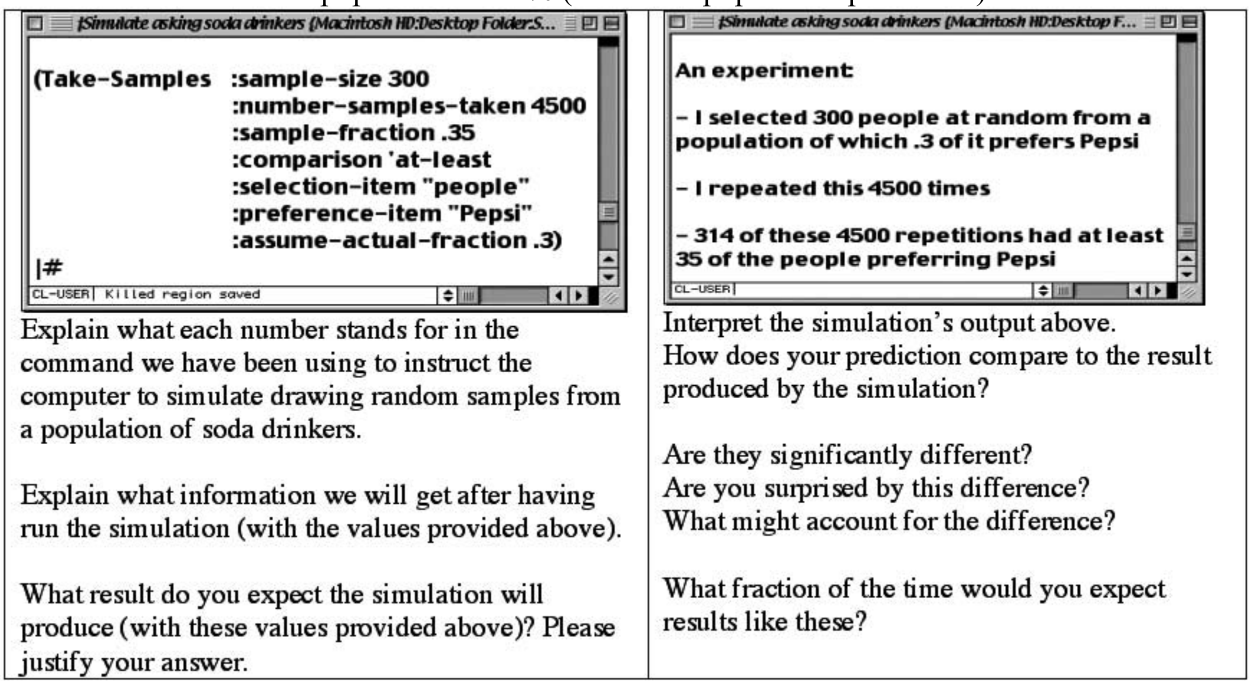


Figure 1. The interface used by Saldanha and Thompson (2002).

Consider the following excerpt of a student named "D":

D: Ok. It's asking…the question is…like "do you like Garth Brooks?". You're gonna go out and ask 30 people, it's[3] gonna ask 30 people 4500 times if they like Garth Brooks. The uh…(talks to himself) what's this? Let's see…the actual…like the amount of people who actually like Garth Brooks are…or 3 out of 10 people actually prefer like Garth Brook's music. (Saldanha & Thompson, 2002, p. 262)

This excerpt demonstrates the action conception of sampling/sampling distribution. The population of familiar objects are people who either liked or disliked Garth Brooks's music. D wanted a count of people who like Brooks's music; the mental process the student had the computer perform was to count the number of people who like Brooks's music out of the random sample of 30 people. The stochastic process that D might have in mind here references what

[3] The "it" that D referred to is the computer simulation program.

Saldanha and Thompson (2002) refer to as the first two phases of the sampling process. D imagined drawing a sample of thirty people from the population, counting those who like Garth Brooks's music (phase one) and then repeating this process 4499 more times (phase two). Thus the action conception of sampling/sampling distribution involves the first two phases of the sampling process in relation to a collection of familiar objects and getting a collection of values for a statistic of interest.

Just as in the case of the action conception of function, the observer hypothesizes that the individual has a vantage point within the sequence for the action conception of sampling/sampling distribution. This means that the individual focuses on the sequence of steps (envision/carry out a stochastic process of drawing a sample from the population of interest, conceive of a statistic of interest, find the value of that statistic) and has to constantly work through each step to arrive at the value of the statistic of interest. Additionally, this inner vantage point indicates that the individual is also constantly looking at the objects/people that comprise the sample in order to arrive at a value of the statistic of interest.

Process Conception of Sampling/Sampling Distribution

As an individual continues to work with phase one and two of the sampling process, the individual maybe begin to anticipate that for any sample taken from the population, a value of the statistic of interest may be obtained. This is to say that the individual may begin to reason that a value for the statistic of interest maybe found without actually having to go through the individual steps of the sequence as laid out in phase one and two of the sampling process. Further, the individual moves into the third phase of the sampling process as described by Saldanha and Thompson (2002); partitioning the collection of values for the statistic to make a judgment. Here the individual creates categories that the value of the statistic from each sample maybe placed into. As the individual creates these categories, he/she begins developing a proto-sampling distribution. This suggests that the individual has an L2 conception of sampling in the MCS framework. If an observer hypothesizes that the individual's understanding of sampling/sampling distribution is best characterized by the individual thinking through the three phases of sampling, anticipating the generation of values of the statistic of interest, and developing a proto-sampling distribution, then the observer would refer to this conception as the *process conception of sampling/sampling distribution*.

The following excerpt provides an example of the process conception of sampling/sampling distribution:

> D: When you go out and take one sample of 30 people, the cut off fraction means that if you're gonna count, you're gonna count that sample, if like 37% of the 30 people preferred Garth Brooks. And then it's going to tally up how many of the samples had 37% people that preferred Garth Brooks. (Saldanha & Thompson, 2002, p. 262)

In particular, notice how D focused not on the sequence of steps defined by phases one and two of the sampling process, but rather categorizing the values of the statistic of interest he anticipates will be produced by phases one and two. D constructed two categories for the sample proportions (people who like Garth Brooks's music); one category for sample proportions above a threshold (the 37%) and one category for sample proportions below the threshold. In creating these two categories, D constructed a proto-sampling distribution.

While there is some sense of sampling variation within the action conception of sampling/sampling distribution, an individual with the process conception of sampling/sampling distribution must realize that variation in the sample values of the statistic is not only possible, but also inevitable. With the action conception, an individual is much more focused on the sequence of steps to get a value of the statistic and is not likely to focus on differences in the values of the statistic. If the individual does look at the differences in values of the sample statistic, he/she may attribute the variation to causal factors rather than chance in the sense of Konold's (1989) outcome approach. However, the variation of the sample values of the statistic of interest is an integral part of the process conception of sampling/sampling distribution. Without the expectation of variation amongst the sample values, the student would not see a need to create categories of those values and thus would not necessarily develop a proto-sampling distribution.

Again, the observer may hypothesize that the individual has a vantage point that is outside of the sequence of steps laid out by phases one and two of the sampling process. When the vantage point is external to the sequence, the individual's attention shifts to the sample values of the statistic per se. Additionally, the individual is not concerned with looking that objects/people that make up the sample, but rather the whole sample. This shift in vantage point supports moving into phase three of the sampling process and the L2 conception of sampling from MCS.

Object Conception of Sampling/Sampling Distribution

As the individual continues working with sequence of steps (the three phases), the individual anticipates the sample values of the statistic of interest and continues to build the proto-sampling distribution. As the individual continues to work with the three phases of the sampling process, the individual may begin to unpack and repack the proto-sampling distribution. This unpacking and repacking of the proto-sampling distribution could lead to the individual to notice characteristics of the proto-sampling distribution. For instance, the individual might see that the sample values of the statistic appear to be heavily clustered around a couple of categories (categories with higher frequencies) and less clustered for other categories (categories with low frequencies). When the characteristics that the individual notices begin to support the individual in making judgments about the expectation of a particular kind of sampling outcome, then the individual may have the L3 conception of the MCS framework. If the observer hypothesizes that an individual's understanding of sampling/sampling distribution is best described by the individual using the three phases of the sampling process to construct the proto-sampling distribution and the individual using characteristics of that proto-sampling distribution to make a judgment about the expectation of a particular outcome, then the observer could refer to this conception as the *object conception of sampling/sampling distribution.*

The following excerpt demonstrates that D is progressing towards the object conception of sampling/sampling distribution; the instructor (I) of the teaching experiment prompts D for some clarification.

D: If like…if you represent—if you give it like the split of the population and then you run it through the how—number of samples or whatever it'll give you the same results as if—because in real life the population like of America actually has a split on whatever, on Pepsi, so it'll give you the same results as if you actually went out, did a survey with people of that split.

I: Ok, now. What do you mean by "same results"? On any particular survey at all—you'll get exactly what it—?

D: No, no. Each sample won't be the same but it's a…it'd be…could be close, closer…

I: What's the "it" that would be close?

D: If you get…if you take a sample…then the uh…the number of the like whatever, the number of "yes's" would be close to the actual population split of what it should be.

I: Are you guaranteed that?

D: You're not guaranteed, but if you do it enough times you can say it's with like 1 or 2% of error depending upon uh how many times—I think—how many times you did it. (Saldanha & Thompson, 2002, pp. 265–266)

At times, distinguishing what exactly D refers to when he says "it" is difficult. However, D is talking about the purpose of sampling. As in the other excerpts, D did know the population parameter, however D still anticipated a plurality of sampling outcomes with some sense of variance in the sample values. There is evidence for this in D's first statement ("...number of samples…" and D's second statement ("No, no. Each sample won't be the same…"). D could be thinking about a proto-sampling distribution and thinking about making judgments with that proto-sampling distribution. There is some evidence to support this in D's third line, D anticipated taking another sample ("…if you take a sample…the number of "yes's" would be close to the actual population split") and then make a judgment about the representativeness a new sampling outcome ("…it's with like 1 or 2% of error…"). What, if any, characteristics that D might see about the proto-sampling distribution are unclear. Thus, an observer might hypothesize that D is progressing towards an object conception but has not actually reached an object conception of sampling/sampling distribution.

As before, there is once again a shift of vantage point of the individual hypothesized by the observer. The vantage point is again outside of the sequence (the phases of sampling), but at a greater distance. This vantage point has the individual now focusing on the proto-sampling distribution and a new sample value of the statistic of interest. This change in vantage point once again supports the L3 conception in the MCS framework.

Other Perspectives and Conclusion

There has been some work done focusing on the (mis-) conceptions that adults have on sampling and other statistical topics (e.g., Kahneman's, Solvic's, and Tversky's 1982 book Judgment Under Uncertainty: Heuristics and Biases) as well as some work on children's conceptions (Lajoie, Jacobs, & Lavigne, 1995) of statistical topics. With reference to children's understanding of sampling, there have been several articles that focus on different aspects of sampling than this proposed framework. Watson and Moritz (2000a, 2000b) looked at a collection of Australian students at various grade levels. Using Watson's statistical literacy framework, they developed six categories for classifying the sampling tendency students gravitated towards (e.g. small samplers without selection). Watson and Mortiz's six categories are based on the size and type of sample that the student constructed. These characterizations could provide an opportunity to return to the assumptions of the APOS-S/SD framework and theorize what might happen when the epistemic individual has a disposition of say a small sampler with pre-selection of results. The role that sample size plays with sampling and sampling distribution is important. While the presented perspective does not address the role that an individual's meanings for sample size plays into his/her understandings of sampling/sampling distribution, this perspective does not disallow for the addition of sample size to the framework.

Jacobs (1997) also looked at children's understanding of sampling centered on how students (fourth and fifth graders) evaluated different sampling methods and how those students drew conclusions based upon presented samples (with sampling methods). Jacobs's work provides a potential starting place for tweaking the assumption in the present framework of the individual's understanding of stochastic process that generates the sample. For instance, if the epistemic individual holds a meaning that a self-selected sampling process is valid, then how we might describe that individual's scheme for Sampling/Sampling Distribution would be quite different from what has been presented.

This framework is related (by necessity) to that of Saldanha and Thompson (2002, 2007) in that this perspective takes a distributional approach to sampling. This is to say that the student's understanding of sampling is placed within his/her ability to keep in mind a sampling distribution. The work of Saldanha and Thompson greatly influenced this framework. While Saldanha and Thompson laid out a three phase sampling process and a multiplicative conception of sampling, this framework serves as an attempt to extend their work. The work of Saldanha and Thompson (2002) described different levels at which a student might need to think multiplicatively but did not address how a student might move from one level to another. APOS-S/SD serves as one framework that provides a stepping-stone towards a developmental framework to extend the MCS and suggest one hypothesis for how an epistemic individual might develop a richer understanding of sampling, sample, and sampling distribution.

The Action, Process, Object, and Schema Theory for Sampling/Sampling Distribution framework hypothesizes an extension of Saldanha and Thompson's (2002) multiplicative conception of sampling. Specifically, this perspective makes use of elements of the APOS theory to hypothesize how an epistemic individual's understanding of sampling and sampling distribution based on the multiplicative conception of sampling might develop under a very strict set of assumptions. Given the set of assumptions, the presented theoretical perspective should not be taken as a developmental trajectory but rather as a first hypothesis on the long road to researching a developmental trajectory.

References

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the Process Conception of Function. *Educational Studies in Mathematics*, *23*(3), 247–285.

Clark, J., Kraut, G., Mathews, D., & Wimbish, J. (2007). *The Fundamental Theorem of Statistics: Classifying Student understanding of basic statistical concepts*. Retrieved from http://www1.hollins.edu/faculty/clarkjm/stat2c.pdf

Dubinsky, E., & McDonald, M. A. (2002). APOS: A constructivist theory of learning in undergraduate mathematics education research. *The teaching and learning of mathematics at university level*, 275–282.

Dubinsky, E., Weller, K., McDonald, M. A., & Brown, A. (2005a). Some historical issues and paradoxes regarding the concept of infinity: An Apos-Based analysis: Part 1. *Educational Studies in Mathematics*, *58*(3), 335–359.

Dubinsky, E., Weller, K., McDonald, M. A., & Brown, A. (2005b). Some historical issues and paradoxes regarding the concept of infinity: An APOS analysis: Part 2. *Educational Studies in Mathematics*, *60*(2), 253–266. doi:10.1007/s10649-005-0473-0

Jacobs, V. R. (1997). Children's Understanding of Sampling in Surveys. (p. 37). Presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. Retrieved from http://www.eric.ed.gov/PDFS/ED411247.pdf

Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction*, *6*, 59–98. doi:10.1207/s1532690xci0601_3

Lajoie, S. P., Jacobs, V. R., & Lavigne, N. C. (1995). Empowering children in the use of statistics. *The Journal of Mathematical Behavior*, *14*(4), 401–425. doi:10.1016/0732-3123(95)90039-X

Lipson, K. (2003). The role of the sampling distribution in understanding statistical inference. *Mathematics Education Research Journal*, *15*(3), 270–287.
Mathews, D., & Clark, J. M. (2007, July 25). *Successful Students' Concpetions of mean, Standard Deviation, and the Central Limit Theorem*. Unpublished. Retrieved from http://www1.hollins.edu/faculty/clarkjm/stats1.pdf
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics□: an overview.* Reston, VA: National Council of Teachers of Mathematics.
National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.
Saldanha, L., & Thompson, P. W. (2002). Conceptions of Sample and Their Relationship to Statistical Inference. *Educational Studies in Mathematics*, *51*(3), 257.
Saldanha, L., & Thompson, P. W. (2007). Exploring connections between sampling distributions and statistical inference: an analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education*, *2*(3), 270–297.
Sfard, A. (1992). Operational origins of mathematical notions and the quandary of reification - the case of function. In E. Dubinsky & G. Harel (Eds.), *The Concept of Function: Aspects of Epistemology and Pedagogy*. Washington, D.C.: Mathematical Association of America.
Sfard, A. (1994). Reification as the Birth of Metaphor. *For the Learning of Mathematics*, *14*(1), pp. 44–55.
Simon, M., Saldanha, L., McClintock, E., Akar, G. K., Watanabe, T., & Zembat, I. O. (2010). A developing approach to studying students' learning through their mathematical activity. *Cognition and Instruction*, *28*(1), 70–112.
Tall, D., & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with Particular Reference to Limits and Continuity. *Educational Studies in Mathematics*, *12*(2), 151–169.
Thompson, P. W. (1994). Students, functions, and the undergraduate curriculum. *Research in collegiate mathematics education*, *1*, 21–44.
Thompson, P. W. (2013, February 21). *Under The Radar: Foundational Meanings That Math Professors (Mistakenly) Assume Teachers Have*. Plenary presented at the 16th Annual Conference on Research in Undergraduate Mathematics Education, Denver.
Watson, J. M., & Moritz, J. B. (2000a). Development of understanding of sampling for statistical literacy. *The Journal of Mathematical Behavior*, *19*(1), 109–136. doi:10.1016/S0732-3123(00)00039-0
Watson, J. M., & Moritz, J. B. (2000b). Developing Concepts of Sampling. *Journal for Research in Mathematics Education*, *31*(1), 44–70.
Wilson, M. (2005). *Constructing measures□: an item response modeling approach*. Mahwah N.J.: Lawrence Erlbaum Associates.

ILLUSTRATING A THEORY OF PEDAGOGICAL CONTENT KNOWLEDGE FOR SECONDARY AND POST-SECONDARY MATHEMATICS INSTRUCTION

Shandy Hauk	Allison Toney	Billy Jackson	Reshmi Nair	Jenq-Jong Tsay
WestEd	U. North Carolina	U. West	U. Northern	U. Texas –
shauk@wested.org	Wilmington	Georgia	Colorado	Pan American

The accepted framing of mathematics pedagogical content knowledge (PCK) as mathematical knowledge for teaching has centered on the question: What mathematical reasoning, insight, understanding, and skills are required for a person to teach elementary mathematics? Many have worked to address this question in K-8 teaching. Yet, there remains a call for examples and theory in the context of teachers with greater mathematical preparation and older students with varied and complex experiences in learning mathematics. In this theory development report we offer background and examples for an extended theory of PCK – as the interplay among conceptually-rich mathematical understandings, experience in and of teaching, and multiple culturally-mediated classroom interactions.

Keywords: Pedagogical content knowledge, Discourse, Intercultural awareness

Since Shulman's (1986) seminal work, a rich collection of theories and measures of mathematics pedagogical content knowledge (PCK) continues to grow (e.g., Ball, Thames, & Phelps, 2008; Hill, Ball, & Schilling, 2008; Silverman & Thompson, 2008). However, the work to date on early grades (K-8) teacher development includes little in the way of the classroom sociology and advanced mathematical understandings such as are found in high school and college. There is a need for examples and theory in the context of teachers with greater mathematical preparation and older students with varied and complex experiences in learning mathematics (Speer & King, 2009).

The framing of knowledge for teaching in the K-8 arena has centered on the question: What mathematical reasoning, insight, understanding, and skills are required for a person to teach elementary mathematics? Many have worked to develop measures to address this question, most notably Ball and colleagues (Hill, Ball, & Schilling, 2008). In their work they have defined three types of subject matter knowledge (SMK) and three types of PCK as the domains of "mathematical knowledge for teaching" (p. 377; see Figure 1). Even with their carefully developed model, challenges exist in identifying and measuring PCK (pp. 396-398).

Speer and King (2009) have offered insight into the different demands of a theory of SMK for secondary and post-secondary instruction. We extend the exploration of differing demands and focus on the PCK half of the picture (see Figure 1, at right): knowledge of curriculum, of content and students (KCS), and of content and teaching (KCT).

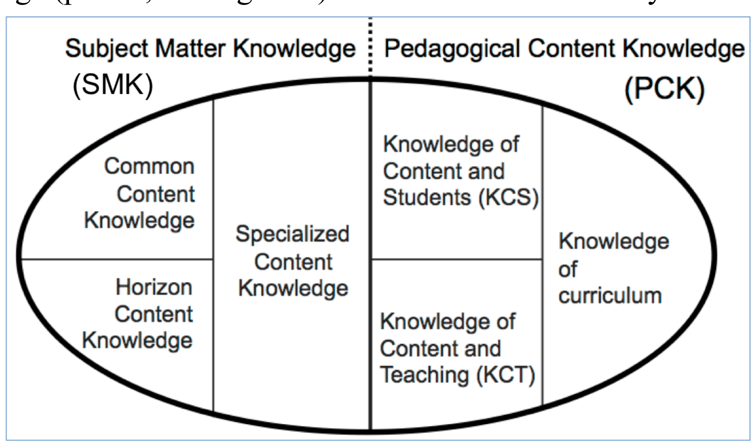


Figure 1. Dimensions of knowledge for teaching from Hill, Ball, & Schilling (2008).

Hill, Ball, & Schilling (2008) acknowledged the problematic nature of identifying types of "knowledge" and have speculated on the need for alternate conceptualizations (e.g., perhaps as "reasoning about") or additional constructs, to capture the multi-dimensional nature of PCK. Other researchers have offered a supplement to the K-8 view, emergent from radical constructivist perspectives (i.e., Piagetian). It is the idea that for some, PCK is "predicated on coherent and generative understandings of the big mathematical ideas that make up the curriculum." (Silverman & Thompson, 2008, p. 502). In this framing, PCK grows when a teacher gets better at the transformation of personal and intimate forms of mathematical knowing. Our purpose in building theory is to describe and illustrate an unpacking of these ideas – attending to people's ways of understanding, thinking, and reasoning about and through mathematics in order to teach, while also attending to the reality of culturally heterogeneous classroom contexts.

Here we report on our efforts to develop an expanded theory and model of PCK that considers a key aspect of Shulman's (1986) original framing that is absent in existing models. Based on work discussed below, it is called *knowledge of discourse*. This brings to PCK the mathematical "syntax" that was part of Shulman's description:

> The syntactic structure of a discipline is the set of ways in which truth or falsehood, validity or invalidity, are established... Teachers must not only be capable of defining for students the accepted truths in a domain. They must also be able to explain why a particular proposition is deemed warranted, why it is worth knowing, and how it relates to other propositions, both within the discipline and without, both in theory and in practice… This will be important in subsequent pedagogical judgments regarding relative curricular emphasis. (Shulman, 1986, p. 9)

Ultimately, we seek to develop theory and measurement tools/guidelines that allow exploration of questions such as: *What is the interplay among mathematical understandings, teaching, and culturally mediated communication in defining and growing PCK?*

Our proposed framework relies on three existing theories related to human interaction in mathematics teaching and learning: for discourse, for intercultural awareness, and for PCK. We start with brief definitions associated with "discourse," make a foray into some key ideas in intercultural orientation, and then describe our model with additional PCK constructs. We conclude with two classroom vignettes and brief analyses of them to illustrate the theorized PCK constructs. These illustrations are *not* definitions. They are offered as anchors for discussion.

Background on d/Discourse

In his review of over one hundred research publications in mathematics education that reported on "discourse," Ryve (2011) concluded that conceptualizations of discourse are varied in detail and diverse in scope. He noted that the field would benefit from explicit definitions for "discourse" each time it is used in reporting research or theory. What Ryve found in common across the reviewed articles was that the conceptions of "discourse" could be understood through the work of Gee (1996), who distinguished between "big D" Discourse and "little d" discourse.

A classroom culture is a set of values, beliefs, behaviors, and norms shared by the teacher and students that can be reshaped by the people in the room (Hammer, 2009). Though not everyone in the classroom may describe the culture in the same way, there would be a general center of agreement about a set of classroom norms, values, beliefs, and behaviors. Whereas Gee's (1996) "little d" discourse is about language-in-use (this may include connected stretches of utterances and other agreed-upon ways of communicating mathematics such as symbolic statements or graphs), Discourse ("big D") includes little d discourse *and* other types of communication that happen in the classroom (e.g., gestures, tone, pitch, volume, and preferred ways of presenting information). The forms of communication in discourse are usually explicit and observable, while the culturally embedded nature of communication in Discourse is largely implicit. Gee's Discourse also includes Shulman's attention to syntax:

> A Discourse is a socially accepted association among ways of using language, other symbolic expressions, and 'artifacts', of thinking, feeling, believing, valuing, and acting that can be used to identify oneself as a member of a socially meaningful group or 'social network', or to signal (that one is playing) a socially meaningful 'role' (p. 131).

That is, as part of PCK, there is knowledge for working effectively with the multiplicity of Discourses students, teacher, curriculum, and school bring into the classroom. Each Discourse includes a cultural context. Discourses may differ from person to person or group to group. The ways that teachers and learners are aware of and respond to multiple cultures is a consequence of their orientation towards cultural difference, their *intercultural orientation*. We come back to intercultural orientation after unpacking what we mean by Discourse a bit more.

The "big D" Discourse of academic mathematics values particular kinds of "little d" discourse. Valued inscriptions are logico-deductive (e.g., proof) and figural (e.g., representations such as graphs of functions or diagrams of relationships or mappings). Especially valued in advanced mathematical discourse are explanation, justification, and validation (Arcavi, Kessel, Meira, & Smith, 1998; DeFranco, 1996; Weber, 2004). As in other fields, instructors ask questions to evaluate what students know and to elicit what students think. For instance, a model of classroom interaction common in the U.S. is the dialogic pattern of *initiation –response – follow-up* or *I•R•F* structure (Mehan, 1979; Wells, 1993). In college classrooms, this is most often initiated by teachers, but not exclusively so, and the (implicit) rules for how initiating, responding, and following-up will happen are worked out by the people in the room (Nickerson & Bowers, 2008). These rules make up one aspect of what Yackel and colleagues have called "socio-mathematical norms" (Yackel, Rasmussen, & King, 2000).

In his ethnographic work, Mehan identified four types of teacher questions (see Table 1). Research suggests that U.S. mathematics instructional practice lives largely to the left of Table 1 (Stigler & Hiebert, 2004; Wood, 1994). The unfortunate aspect here is not the fact that evaluative questions are common but that the eliciting questions, in the right column, are not. These more complex spurs for discourse can lead to iterative patterns that cycle through and revisit the frame of reference "in ways that situate it in a larger context of mathematical concepts" and foster "mathematical meaning- making" (Truxaw & DeFranco, 2008, p. 514). The use of process and metaprocess questions, for example as *follow-up (F)*, readily expands discourse into the "reflective toss" realm of comparing and contrasting different ways of thinking (with justification but without judgment), monitoring of a discussion itself, as well as attending to the evolution of the thinking of others and self (van Zee & Minstrell, 1997).

Table 1. Initiate-Respond-Follow-up (*I•R•F*) question types and anticipated response types.

Evaluate what students know	***Elicit what students think***
Choices – response constrained to agreeing or not with a statement (e.g., Did you get 21?) ***Products*** – response is a fact (e.g., What did you get?)	***Processes*** – response is an interpretation or opinion (e.g., Why does 21 make sense here?) ***Metaprocesses*** – response involves reflection on connecting question, context, and response (e.g., What does the 21 represent? How do you know?)

Another important aspect of Discourse is in interaction for teaching. Piaget identified assimilation and accommodation as two interactive processes to explain an individual's adaptation to achieve cognitive equilibration and learn (Driscoll, 1994). Humans are pattern-seekers looking for patterns to recognize for assimilation. If assimilation fails, people may create their own interpretation of ideas, based on available perceptions, for accommodation. From this perspective, teaching is the act of providing productive cognitive conflict so that learners may accommodate their existing schemes, iteratively, in ways that incorporate rigorous mathematical schemes. That is, concept images are challenged repeatedly by cognitive disequilibration to foster the development of the associated concept definition (Tall & Vinner, 1981). This is in contrast to pseudo-assimilation (e.g., about function; Zandieh, 2000).

As an example, consider the types of pseudo-assimilation discussed by Bair and Mooney (2013). They offered examples of problematic instruction on the distributive property such as "FOIL" and "bam-baming" two negative signs to a positive in an expression like $4-(6-3x)$. Although aiming to reduce what learners may find cognitively overwhelming, these may lead students to unproductive generalizations and counter-productive decisions about mathematical meanings. Similarly, Temple and Doerr (2012) note the importance of developing fluency in the mathematical register – thought and speech inform each other and using technical vocabulary can support mathematical meaning-making. Discourse is central in our effort to bring to PCK theory an explicit attention to the use of language and the dense set of values about mathematical appropriateness, clarity, and precision that are integral to thinking, learning, and communicating mathematics. In what follows, we use d/Discourse in Gee's (1996) culturally informed way.

Intercultural Orientation

The construct of "big D" Discourse as part of mathematics PCK pivots on the idea of intercultural orientation. Our referent framework is the *Developmental Model of Intercultural Sensitivity* (Bennett & Bennett, 2004). The developmental continuum of orientations towards awareness of cultural difference, of "other," 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 first move, from denial to the "polarization" orientation, comes with the recognition of difference, of light and dark in viewing a situation (e.g., Figure 2a).

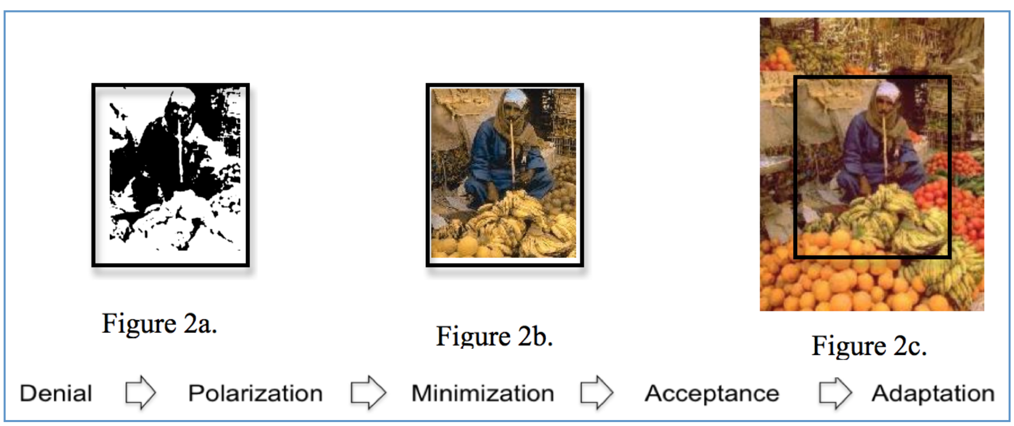


Figure 2. Intercultural orientations and developmental continuum.

The polarization orientation is driven by the assumption "Everybody should be like me/my group" and is an orientation that views difference in terms of a stark "us" and "them." Evaluative prompts about student thinking (left side of Table 1) are more likely for this orientation. Moving along the continuum towards ethno-relative perspectives leads to a minimizing of difference, focusing on similarities, commonality, and presumed universals (e.g., biological similarities – we all have human brains so we all learn math essentially the same way; and values – we all know the difference between right and wrong and naturally will seek right). This is the "minimization" orientation. A person with this orientation will be blind to recognition and appreciation of subtleties in difference (e.g., Figure 2b, a representation of, literally, the view of a colorblind person). The minimization orientation tends to take the form of ignoring fine detail in how people might have differing ways of thinking. For example, efforts at eliciting d/Discourse (right side of Table 1) may take the form of *listening for* particular ways of thinking. Transition from a minimization orientation to the "acceptance" orientation involves attention to nuance and a growing awareness of self and others as having culture and belonging to cultures (plural) that may differ in both obvious and subtle ways. While aware of difference and the importance of relative context, how to respond and what to respond *in the moment of interaction* is still elusive. From this orientation, classroom d/Discourse may include process and metaprocess prompts, but sustained cycles of such interactions can be challenging to maintain in the immediacy of dynamic classroom conversation. The transition to "adaptation" involves developing frameworks for perception, and responsive skills, that attend to a spectrum of detail in an interaction (e.g., the detailed and contextualized view in Figure 2c). Adaptation is an orientation where one is aware of multiple relative perspectives, and may – without violating one's authentic self – adjust communication and behavior in contextually appropriate ways. There is an instrument for measuring general intercultural orientation (see idiinventory.com). The central idea here is that such orientations are learned, developmentally (Bennett, 1993, 2004; DeJaeghere & Cao, 2009).

Extended Model of Pedagogical Content Knowledge

While Hill, Ball, and colleagues took a classical measure theory approach to identifying and assessing teacher knowledge, we continue to investigate a non-linear alternative (i.e., instead of the traditional linear methods such as hierarchical linear modeling). In particular, our current approach focuses on PCK in terms of four areas of professional understanding: Knowledge of Discourse, Curricular Thinking, Anticipatory Thinking, and Implementation Thinking. These four areas connect in many ways with the Knowledge of Curriculum, Knowledge of Content and Students (KCS), and Knowledge of Content and Teaching (KCT) from Figure 1. They differ, however, in that each is a kind of proceptual understanding (Gray & Tall, 1994), with thinking that integrates relational components along with instrumental ingredients (Skemp, 1976). We seek to identify, prompt for, and assess the connected and overlapping relational aspects, especially in how the three types of thinking (curricular content, anticipatory, implementation) interact with knowledge of curriculum, KCS and KCT to be generated by and generative of Knowledge of Discourse.

Hill, et al. (2008) acknowledge the importance of teacher knowledge of standard and non-standard mathematical representations and communication, but knowledge of d/Discourse as we construe it – composed of discourse and Discourse – does not appear explicitly in their model. One way of visualizing our extension, that highlights and focuses on the interplay among the components of the new and existing models, is as the surface of a tetrahedron whose base is the existing model with a new vertex of Knowledge of Discourse (see Figure 3). We have focused on knowledge of discourse and the three "edges" connecting it to the components in Figure 3 (Hauk, Jackson, & Noblet, 2010). These edges are labeled as "ways of thinking" in the sense of Harel (2008). We continue to explore the possibility of the "knowledge of" areas being taken as "(ways of) understanding" (Harel, 2008).

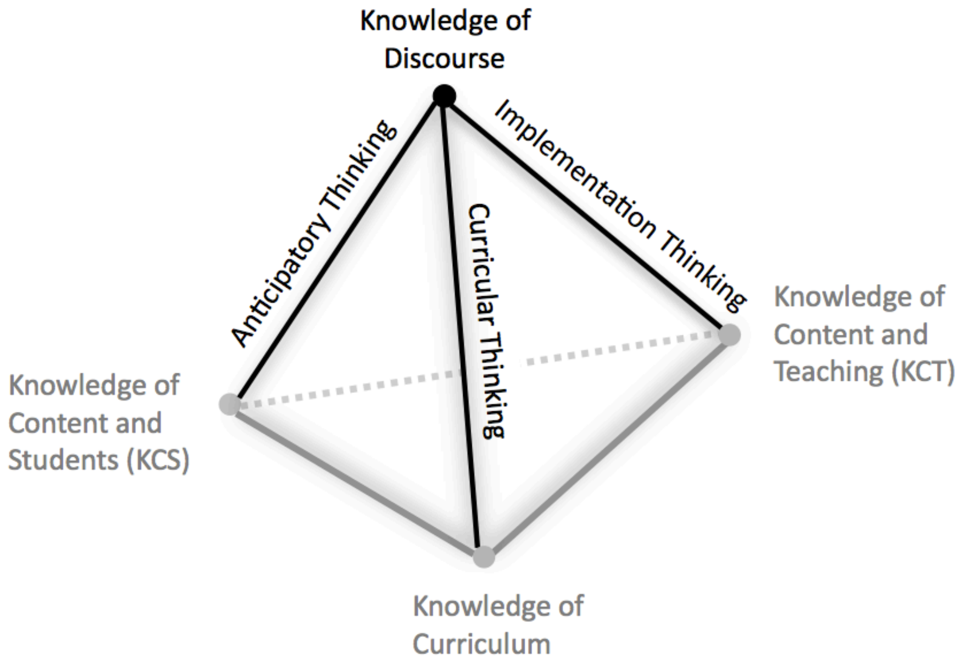


Figure 3. Tetrahedron - vertices, edges, and surfaces - as a way to visualize PCK components and relationships. Corners of the base are PCK dimensions from Figure 1.

Knowledge of Discourse is d/Discourse knowledge about the culturally embedded nature of inquiry and forms of communication in mathematics (both in and out of educational settings).

Curricular Thinking is ways of thinking about mathematical topics, procedures, and concepts as well as the relationships among them, and conventions for reading, writing, and speaking them, in curricula. In its most robust form, this part of PCK contributes to what Ma (1999) called "profound understanding of mathematics" (p. 120). In combination, curricular content and d/Discourse are the home of Simon's (2006) "key developmental understandings."

Anticipatory Thinking is ways of thinking about (strategies, approaches to) how learners may engage with content, processes, and concepts. It includes awareness of and responsiveness to student thinking. Part of anticipatory development involves what Piaget called "decentering" – building skill in shifting from an ego-centric to an ego-relative view for seeing or communicating about an idea or way of thinking from the perspective of another (e.g., eliciting, noticing, and responding to student thinking; Carlson, Moore, Bowling, & Ortiz, 2007). Teachers with complex anticipatory thinking manage the tensions among their own instrumental and relational understandings of mathematics and its learning and those of their students (Skemp, 1976). Such perspective-shifting is deeply connected to d/Discourse through the awareness of "other" as different from "self." We see this as intimately connected to intercultural orientation.

Implementation Thinking. This is ways of thinking about (strategies, approaches to) how to enact teaching intentions in the classroom. Moreover, for us, it includes how to adapt teaching according to content and socio-cultural context and act on decisions informed by d/Discourse as well as curricular content and anticipatory ways of thinking. We do not argue for an intention to enculturate in the sense of Kirshner's (2002) "teaching as enculturation" (i.e., to identify a reference culture and then target instruction for students to acquire particular dispositions). Nor do we propose his alternate framings (habituation, construction) or any other preference for a particular implementation paradigm.

Vignettes and Discussion

Over the last 10 years, the authors have been involved in a variety of ways in research and professional development with post-secondary faculty, in-service secondary mathematics teachers, and their students. In that work, mathematically trained stakeholders regularly ask us for examples and non-examples of PCK in use. The two vignettes included here are about Teacher Pat, a mathematics doctoral student at a large university whose teaching assistantship includes leading a recitation section in undergraduate abstract algebra. The examples are based on real classroom transcripts from various research projects by colleagues and ourselves. Vignette 1 is Teacher Pat in the first year of teaching a group theory recitation.

Vignette 1 – Snapshot of an Abstract Algebra Classroom

Pat stands in front of the whiteboard. Twenty students are seated among 36 small desk-chairs arranged in 6 rows of 6, all facing the front of the room. Problem on the board reads: *Let G be a group of even order. Prove or disprove that there exists a nonidentity element x of G such that* $x^2 = e$. Shuffling of paper, scratching of pencils, but no voices as students work.

Pat: Okay. Let's talk about this problem a bit. We've been talking lately about planning out our proofs – to decide if we believe the statement to be true before proving it. How is it that you thought about this proof?

Lee: I made a set with an even number of elements and then wrote the inverses. One to two, two to three, and so on.

Pat: Okay *(pause)*, what did you do when you got to the end of your set?

Lee: That went back to one.

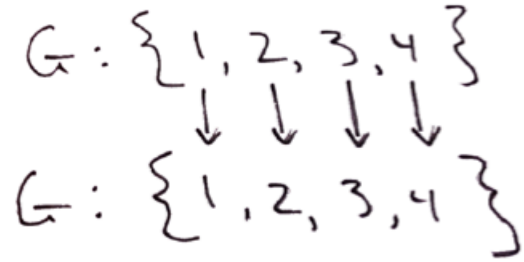
Pat: So, you disproved the statement using a counter example?

Lee (appearing puzzled by the question; looks down at own work; looks back at the problem on the board): Um-

Pat: But this statement is actually a true statement, right? *(To the class)* Right?

(Pat writes "G: {1, 2, 3, 4}" twice on the board, vertically stacked; draws an arrow from 1 on the top line to 1 on the bottom line, and then draws similar arrows for 2, 3, and 4.)

Pat: Here's a set, G, of even order – it has an even number of elements. One goes to one, two to two, three to three, four to four. Each element is its own inverse.



So, we see that the statement can be true.

Lee: Oh! Or can three and four can be each other's inverses and one and two still each be their own inverse?

Pat: Right. As long as the group has an even order, there is always at least one element, in addition to the identity, that is its own inverse. *(Pat looks around the room)* Any questions? Let's write out the proof. *(Turns and erases the board).*

Jackie (quietly, to self) But why couldn't it go around a circle?

Pat: What's that? *(Turns to face the room)*

Jackie: Couldn't one go to two, two to three, three to four, and four back to one?

Pat: No. Then the inverse of the identity wouldn't be the identity. The inverse of the identity always has to be the identity.

Jackie (shrugging): Um, okay.

Pat: Okay. Does anyone have any more questions? Okay. Now that we've planned out our proof, let's write it up formally.

Figure 3. Vignette representing Teacher Pat's instruction in first year of teaching.

Vignette 2 is Pat teaching the same kind of section, after two years that included observing others' classes and participating in seminars about noticing and responding to student thinking.

Vignette 2 – Snapshot of an Abstract Algebra Classroom

Pat stands in front of the whiteboard. Twenty students are seated in small desk-chairs scattered around the room; some are in rows facing the front, while others are clustered facing each other in groups of 2-3. Problem on the board reads: *Let G be a group of even order. Prove or disprove that there exists a nonidentity element x of G such that* $x^2 = e$. Shuffling of paper, scratching of pencils, rumble of voices as students work and talk to each other.

Pat: Let's talk about this problem a bit. From what I could tell walking around the room, some of you found this proof challenging but doable, and some of you found it challenging and struggled to come to a proof. *All* of you found it time consuming *(students nod and a few look at each other and chuckle)*. We've spent some time recently talking about planning out your proofs – to decide if you even believe the statement before tackling it. Let's take a look at somebody's work who feels like they planned out their proof. Could I get a volunteer? Or volunteers, for those of you who worked together?

(Students look around the room at one another; Sam and Lia, whose desks are turned to face each other, make eye contact, Sam raises his eyebrows questioningly, and Lia nods back)

Sam: We'll do it. *(Pat steps aside and gestures to the whiteboard; Sam and Lia come to the board; everyone moves to see the whiteboard)*

Sam (picking up a marker and writing on the board): We started by letting G be this set. *(Sam writes, "G: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}")*

Pat: Okay wait. Can you say why you wrote G as this set of elements? The statement doesn't say anything about a specific set.

Lia: Well, we know G is a group of even order. That means G has an even number of elements. So, as part of our plan we decided to suppose G is this set *(Lia points to the set written on the board)*.

Pat: Do you all agree with Lia? *(Students nod or say "yes")*

Sam (writing the same line "G: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}" below the first line): We can write the inverses below, *(Sam draws an arrow from 1 on the top to 2 on the bottom)*, where one goes to two. And then we can keep going like this: two to three, three to four, and so on *(Sam adds arrows connecting n on top to n+1 on bottom)*…

So, this means that the statement is true and we think we're ready to start proving.

Pat (to the class): Are there any quest- *(Pat interrupts himself)*- I mean, *what* questions do you have for Sam and Lia?

Jonah: Could you say what the arrows mean?

Lia (pointing to Sam's lists on the board): One goes to two, two goes to three, and so on.

Pat: Yes, but could you please explain what "goes to" *means?*

Sam: Inverses.

Dana: So, one is the inverse of two and two is the inverse of three?

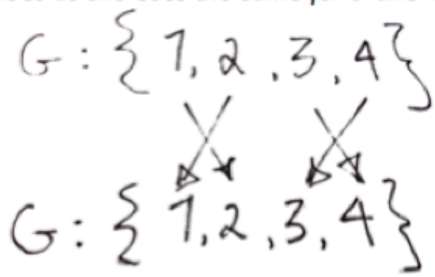
Pat (to class): Do you all agree?

Jonah: No, because then two's inverse would be one.

Pat (to the class) Is that problematic? Think about this for a minute.

(Rumble of voices, shuffling of paper, scratching of pencils; Sam and Lia turn and talk to one another; Lia erases and writes "G: {1, 2, 3, 4}" twice, vertically stacked, and draws arrows

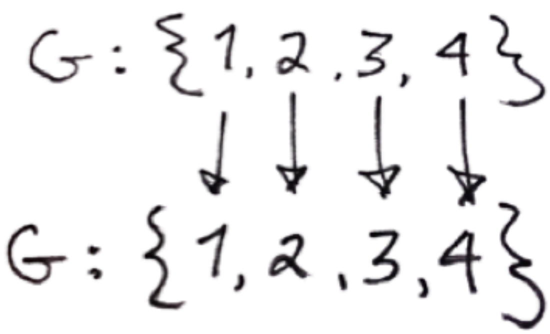
crossing from 1 on top to 2 on the bottom and from 2 in the top line to 1 in the bottom; Sam nods as she does the same for 3 and 4)



Sam (to Pat): We just disproved the statement using a counter example.

Pat (to the class): Let's turn our attention back to Sam and Lia. They say they've found a counter example. *(to Sam and Lia)* Okay, so which element was the identity?

Lia: Well…uh… *(Pause)* The inverse of the identity is the identity!
(She quickly erases the lines and draws a new arrow from 1 on the top line to 1 on the bottom line denoting 1 as the identity)



Lia: Each of the four elements is its own inverse. *(She draws arrows for 2, 3, and 4)*

Nasir: Wait!

Pat: What it is Nasir?

Nasir: (hesitatingly) Can't three and four still be each other's inverses?

Pat (looking around): What do others think?

Dana: Yes! And then two would have to be its own inverse!

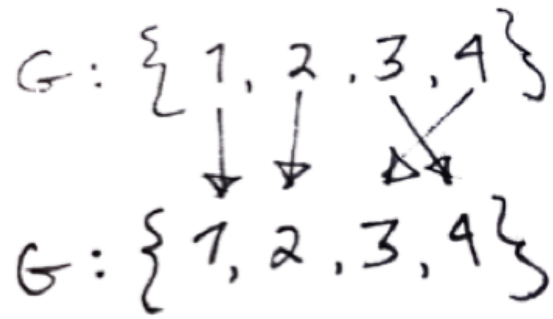
Pat: Slow down. Can you explain Dana?

Dana: Three and four can be each other's inverses. We know one is the identity and it has to be its own inverse. So, two would also have to be its own inverse.

Bo (to Dana): Why would two have to be its own inverse?

Dana (turning to look at Bo): Well, because the order of the group is even.

Lia: So, what you're saying is it could also look like this. *(Lia turns to the board, erases the arrows from 3 and 4 in the top and redraws, crossing them.)*



Lia: Therefore, as long as a group had an even number of elements like in here *(Lia adds a vertical dotted line slicing between 2 and 3 in both sets)*, there would always be *at least* one element in addition to the identity that had to be its own inverse.

Pat: Okay. Thanks, Lia. Thanks, Sam. You two can go ahead and sit down. Did everyone follow that? Everyone take three minutes now to revisit what you have written, and when you're ready we'll get a draft of the formal proof on the board.

Figure 4. Vignette of Teacher Pat's classroom instruction in third year of teaching.

As part of the responsibilities for leading a recitation section, the lead professor with whom Pat worked, Dr. Gold, required graduate students to sit in on the main class meetings. For the interested reader we also offer an online appendix (Toney, Hauk, & Hsu, 2013) – a snapshot of Professor Gold's classroom on the first day of the semester. The idea of the appendix vignette is that Professor Gold works from the first day to establish classroom social and sociomathematical norms that are dialogic. Gold's voice is not the only voice in the room and Gold works to have students feel comfortable with their voices being heard (e.g., orchestrating whole class discussion by asking for raised hands, pausing, requesting think time, and then asking for hands again). Additionally, the professor's comfort with the curriculum is such that Gold effectively anticipates student struggles with group theory concepts. Starting on the first day, instructional materials support student-centered development of concepts. By the time of Vignette 2, Teacher Pat has spent several years observing Gold, someone whose classroom practice is aligned with the right side of Table 1. The reader is encouraged to read through both Pat vignettes before going on to the discussion below. Reading the appendix vignette from Dr. Gold's class may give helpful information on the kind of instruction that Pat observed.

Knowledge of Discourse. In Vignette 1, Teacher Pat foregrounds the correctness of a way of thinking about mapping out a proof and a single path to that proof. That is, the primary discourse (little "d") in the classroom is largely univocal: Pat's utterances to identify a correct proving procedure. Discourse (big D) is also centered with the teacher, as the explanations valued in the classroom are Pat's. In Vignette 2, Pat repeatedly asks students to explain their thinking and has established an atmosphere where students give and ask for explanations. The utterances in the room are more dialogic than in Vignette 1. To participate in discourse (little "d"), responding students have been asked to offer their own thinking to provide a convincing argument. Eliciting questions by Pat are much more in evidence in Vignette 2. An aspect of the classroom Discourse, then, is that engaging in deep explanation is an expectation of all. The request for and use of student-generated figures on the board is part of the mathematical Discourse as much as the valued behavior of students convincing themselves before presenting a "correct" proof. Common to both vignettes is the use of "goes to" as informal language for "maps to." This is an informal phrase widely used and accepted in advanced mathematics discourse. As noted in the section on d/Discourse, use of informal language constrains and supports learning. It can enhance retention, but may also undermine conceptual accommodation. Students' unchecked use of "goes to" as acceptable mathematical language could result in mathematical inaccuracies. In Vignette 1, Student Lee begins with an abbreviated version of "goes to," which Pat reinforces by using the informal phrase a moment later. While Pat may have a proceptual understanding of mappings, the students may still be juggling process and concept as separate, dis-integrated, mental structures. Use of the phrase may funnel all into the process and bypass connected schematization of "mapping elements to their inverses." In Vignette 2, Pat has the students pause and clarify their meaning to some extent, but a firm disequilibration and clear resolution for students is not portrayed in the piece of class we see in the vignette. In Vignette 1, Pat implements *choice* and *product* questions. If these questions dominate a teacher's contributions to discourse, then multiple disconnected *I•R•F* interactions can yield a teacher-regulated kind of interaction that does not include deep participation by students. This can be true even in inquiry-based instruction (Nassaji & Wells, 2000; Wertsch, 1998).

Curricular Thinking. There are subtle and distinct differences between the two vignettes with respect to Pat's content questioning. In Vignette 1, Pat's responses include immediate correct or incorrect feedback. Pat also mentions briefly the idea of a larger goal of planning a proof, while an integration of underlying rationales for such planning is implicit. Unlike Vignette 1, in Vignette 2, Pat's questioning provides cognitive conflict about central concepts (identity, inverse, and to some extent, mapping). To resolve the dissonance, students attend to the properties of identity and inverse, and also notice their interaction (the mapping). A potential connection to the next curricular step lurks in the background as Pat ends the segment by directing students to reflect on what they think and opens the door to connecting it to proof.

Anticipatory Thinking. In Vignette 1, Pat demonstrates anticipatory thinking (and *I•R•F* evaluative approval) of a correct proof path expressed as procedural knowledge. Additionally, Pat does not appear to anticipate the variation in student thinking in the room. Based on Lee's explanation, Pat asserts a misconception for Lee (though Lee seems unaware of it). A moment later, Pat evaluates Jackie's statement rather than taking up the statement as an anticipatory opportunity about student confusion. That is, in Vignette 1, Teacher Pat does not appear to anticipate common student struggles, while also noting a (possible) struggle in a way that is not especially productive. This leads to a question about the nature of anticipatory thinking and its relationship to what actually happens in the classroom (i.e., how might anticipatory thinking be seen as subtly and grossly different from implementation thinking). As we see in Vignette 2, anticipation can be a valuable resource for enhancing students' understanding of mathematics. In Vignette 2, Pat asks guiding questions that involve student thinking. Also, Teacher Pat anticipates that students may possess some knowing of properties of identity and inverse, but may not recognize their interaction in the context of the particular proposition in question. Pat looks to elicit an intellectual need for accommodation by having students display and consider the potential mismatch of information through the representations they draw. In Vignette 2, anticipating and eliciting of student thinking are central and are leveraged by Pat as an implementation strategy: students make sense of their rationales as part of proof planning.

Implementation Thinking. Vignette 1 indicates Pat has a proof map in mind to guide steps of an example and Pat's implementation thinking includes putting Pat's idea of a correct solution path into the air in the room. While Teacher Pat's own subject matter knowledge may operate WLOG (without loss of generality), that strategy may not be familiar to or understood by students. That is, Vignette 1 is pedagogy for proceduralizing proof writing (e.g., "What did you get when you got to the end of your set?"). There is no student-to-student interaction and when Pat overhears Jackie's question, the response is to evaluate and correct (staying to the left of Table 1). In Vignette 2, Pat actively elicits and connects student thinking to procedures *and* concepts. Pat's implementation encourages students to make sense of each other's ideas. As students present their ideas, Pat emphasizes reasoning rather than the product (e.g., "Can you say why you write G as this set of elements?"). Pat also uses multiple modes of discourse, including student generated representations and confirming questions in order to support the needs of various students. Teacher Pat asks the students to clarify their terminology and language so others can make sense of it and share their understanding (e.g., "Could you explain what 'goes to' means?"). The connection between what has been put on the board and what "writing a proof" means remains unspecified at the end of Vignette 2. Pat's implementation thinking in Vignette 1 focused on getting the right answer in the air whereas Pat's implementation approach in Vignette 2 seems to incorporate aiming for the next curricular step, attending to student thinking, and building effective discourse through attention to making sense of and reasoning about the mathematics at hand.

Conclusion

Researchers have suggested that some forms of effective teaching may be comparable to improvisational performance (Borko & Livingston, 1989; Bourdieu – see Grenfell & James, 1998; Yinger, 1987). Teaching requires complex management of instructional resources, including the teacher's own subject matter and pedagogical content knowledge. How communication is initiated, normed, and revised in the classroom is shaped by intercultural awareness. We have attempted to capture and include the shaping of classroom mathematics communication in an extended theory of PCK as *Knowledge of Discourse*. It is likely that rich knowledge of discourse would be fundamental to the kind of teaching that can be characterized as effective improvisation. Success is not just about what is said, but also how it is said, as well as the intimacy established among the participants in an improvisational interaction. In the mathematics classroom, teaching extends beyond precise and accurate transmission of facts or uptake by students of information. Rather, it includes taking into account the background and experiences (mathematical and otherwise) of the people in the room, and making decisions informed by that knowledge and instructional context to shape opportunities for learning.

An area of ongoing work for us is the relationship between intercultural orientation and what orientation(s) may be necessary, if not sufficient, for rich d/Discourse development for teaching. In particular, we continue to explore the extant literature on the concept of "decentering" as one potential instantiation of the developmental intercultural continuum that might be seen at work in classrooms. Moreover, the visualization of the extended theory as the vertices, edges, and faces of a tetrahedron may offer a way of articulating how intercultural orientation, as part of d/Discourse, may be seen (tacitly or overtly) in looking at PCK. That is, suppose each of the four faces in Figure 3 represents a multi-dimensional interaction. For example, consider the face at the back of Figure 3; if we label the "edge" between KCT and KCS (perhaps call it *balancing intended and achieved concepts*) then – if we can go this far without breaking the usefulness of the visual model – how might instructional activity near the lower edge of the face be different from instructional activity on the same face, but closer to the Knowledge of Discourse vertex? Perhaps the difference is the nature of decentering. Or, perhaps it is a more complex intercultural constellation of which decentering is part. Conversely, in comparing Vignettes 1 and 2, where might we point or trace a path on the tetrahedron to indicate that Pat built skill in generating and sustaining conceptually focused discourse during instruction?

While the vignettes included here were for a relatively novice college instructor, at the RUME 2013 meeting, presenters and audience members also talked about the situation where a professor is instructor to a room full of in-service secondary mathematics teachers. In our research with a set of mathematics PhD faculty, the distribution of orientations across the developmental continuum pictured in Figure 2 has been centered in minimization with small variance. At the same time, though distributed more widely across the developmental continuum, the in-service teachers in our work (over 100) have intercultural orientations centered at polarization (Hauk, Yestness, & Novak, 2011; faculty and teachers completed measures of intercultural orientation). We have seen many in-service teachers ready to think about and pay attention to how others' approaches to learning might differ from their own. Meanwhile, their professors have tended to minimize difference. So, when teacher-learners spoke in class about their mathematical understandings and how they differed, professors suggested it was most important to see how the approaches were essentially the same. That is, a challenge for the professors was how to notice *nuances* in the differences across teacher-learners' ways of thinking and use that information in their own anticipatory and implementation thinking. Faculty whose instruction of undergraduates looked like Pat in Vignette 2, were more like Pat in Vignette 1 when working with in-service teachers. The diversity of background and content knowledge is much greater in the teacher-learner population than is typical among undergraduate math majors. It may be that the intercultural pressures on Knowledge of Discourse can be so large as to impede flow along the anticipatory and implementation thinking edges of the tetrahedron. A complementary area for research that might illuminate the relationships is looking at the classroom interactions for polarization-centered in-service secondary teachers. A polarization orientation means identifying difference is a ready skill, but identifying and building on commonality is a challenge. We continue to explore what it means to have rich Knowledge of Discourse and how it and orientation towards cultural difference can support teaching that balances and engages with myriad cultures in-the-moment to scaffold effective mathematical communication among all in the room.

Finally, two suggestions for our ongoing work arose out of the lively discussion at the RUME 2013 session. One was the recommendation that development of the theory presented here pursue the distinction between "ways of thinking" and "ways of understanding." Also arising in discussion at the meeting was the suggestion that we consider a further extension of the visual model with the addition of another tetrahedron for SMK, linked to the PCK model through the Knowledge of Discourse vertex. It is still an open question whether this linking could be useful in thinking about, describing, and developing the knowledge used for teaching mathematics.

Acknowledgement

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. DUE0832026. Any opinions, findings and conclusions or recommendations expressed this material are those of the authors and do not necessarily reflect the views of the NSF.

References

Arcavi, A., Kessel, K., Meira, L., Smith, J. P. (1998). Teaching mathematical problem solving: An analysis of an emergent classroom community. In J. Kaput, A. H. Schoenfeld, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 1–70). Providence, RI: American Mathematical Society.

Bair, S. L., & Mooney, E. S. (2013). Mario, monkey man, fish, and bam-bam: Creative terminology in today's classrooms. *The Mathematics Teacher*, *106*, 408-409.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59,* 389-407.

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., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American educational research journal*, *26*(4), 473-498.

Carlson, M. P., Moore, K., Bowling, S., & Ortiz, A. (2007). The role of the facilitator in promoting meaningful discourse among professional learning communities of secondary mathematics and science teachers. In T. Lamberg & L. R. Wiest (Eds.), *Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 841-848). University of Nevada, Reno.

DeFranco, T. C., (1996). A perspective on mathematical problem solving based on the performances of Ph.D. mathematicians. In Kaput, J., A. Schoenfeld, & E. Dubinsky (Eds.), *Research in collegiate mathematics education. II* (pp. 195–213). Providence, RI: American Mathematical Association.
DeJaeghere, J. G., & Cao, Y. (2009). Developing U.S. teachers' intercultural competence: Does professional development matter? *International Journal of Intercultural Relations*, *33*, 437-337.
Driscoll, M. P. (1994). *Psychology of learning for instruction*. MA: Needham Heights.
Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A proceptual view of simple arithmetic. *Journal for Research in Mathematics Education*, *25*, 116-140.
Grenfell, M., & James, D. (1998). Bourdieu and education: Acts of practical theory. London: Falmer.
Grossman, P. L., Wilson, S. M., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. In M. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 23-36). UK: Pergamon.
Hammer, M. (2009). The Intercultural Development Inventory: An approach for assessing and building intercultural competence. In M. A. Moodian (Ed.), *Contemporary leadership and intercultural competence* (pp. 203-217). Thousand Oaks, CA: Sage.
Harel, G. (2008). What is mathematics? A pedagogical answer to a philosophical question. In B. Gold & R. A. Simons (Eds). *Proof and other dilemmas: Mathematics and philosophy* (pp. 265-290). Washington, DC: Mathematical Association of America.
Hauk, S., Jackson, B., & Noblet, K. (2010). No teacher left behind: Assessment of secondary mathematics teachers' pedagogical content knowledge. In S. Brown (Ed.), *Proceedings for the 13th Conference on Research in Undergraduate Mathematics Education (electronic).* PDF available at sigmaa.maa.org/rume/crume2010/Archive/HaukNTLB2010_LONG.pdf
Hauk, S., Yestness, N., & Novak, J. (2011). Transitioning from cultural diversity to intercultural competence in mathematics instruction. In S. Brown (Ed.), *Proceedings for the 14th Conference on Research in Undergraduate Mathematics Education (electronic).*
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,* 372-400.
Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An exploratory study. *Cognition and Instruction*, *26*, 430-511.
Kirshner, D. (2002). Untangling teachers' diverse aspirations for student learning: A cross disciplinary strategy for relating psychological theory to pedagogical practice. *Journal for Research in Mathematics Education, 33*, 46-58.
Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.
Nassaji, H., & Wells, G. (2000). What's the use of 'triadic dialogue'? An investigation of teacher–student interaction. *Applied Linguistics, 21*(3), 376–406.
Nickerson, S. & Bowers, J. (2008). Examining interaction patterns in college-level mathematics classes: A case study. In M. Carlson and C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics education* (pp. 179–190). Washington, DC: Mathematical Association of America.

Ryve, A. (2011). Discourse research in mathematics education: A critical evaluation of 108 journal articles. *Journal for Research in Mathematics Education*, *42*(2), 167-198.

Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education, 11*, 499-511.

Simon, M. A. (2006). Key developmental understandings in mathematics: A direction for investigating and establishing learning goals. *Mathematical Thinking and Learning, 8*(4), 359-371.

Skemp, R. (1976). Relational and instrumental understanding. *Mathematics Teaching, 77,* 20-26.

Stigler, J. & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 12-19.

Temple, C., & Doerr, H. M. (2012). Developing fluency in the mathematical register through conversation in a tenth-grade classroom. *Educational Studies in Mathematics*, *81*(3), 1-20.

Thompson, A. G., Philipp, R. A., Thompson, P. W., & Boyd, B. A. (1994). Calculational and conceptual orientations in teaching mathematics. In A. Coxford (Ed.), *1994 Yearbook of the NCTM* (pp. 79-92). Reston, VA: NCTM.

Toney, A., Hauk, S., & Hsu, E. (2013). Appendix A: Professor Gold and the first day of abstract algebra. Available online at: http://www.bookwalk.org/hauk/RUME2013/PCK/AppendixA.pdf

Truxaw, M. P., &Defranco, T. C. (2008). Mapping mathematics classroom discourse and its implications for models of teaching. *Journal for Research in Mathematics Education*, *39*(5), 489–525.

van Zee, E. & Minstrell, J. (1997). Using questioning to guide student thinking. *The Journal of the Learning Sciences, 6*(2), 227–269.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of professors' lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior, 23*(2), 115–133.

Wells, G. (1993). Reevaluating the IRF sequence: A proposal for the articulation of theories of activity and discourse for the analysis of teaching and learning in the classroom. *Linguistics in Education, 5*(1),1–37.

Wertsch, J. V. (1998). *Mind as action*. New York: Oxford University Press.

Wood, T. (1994). Patterns of interaction and the culture of mathematics classrooms. In S. Lerman (Ed.), *The culture of the mathematics classroom* (pp. 149–168). Dordrecht, The Netherlands: Kluwer.

Yackel, E., Rasmussen, C., & King, K. (2000). Social and sociomathematical norms in an advanced undergraduate mathematics course. *Journal of Mathematical Behavior, 19*(3), 275–287.

Yinger, R. J. (1987, April). *By the seat of your pants: An inquiry into improvisation and teaching.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

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).

DEVELOPING FACILITY WITH SETS OF OUTCOMES BY SOLVING SMALLER, SIMILAR COUNTING PROBLEMS

Elise Lockwood
University of Wisconsin - Madison

Combinatorial enumeration has a variety of important applications, but there is much evidence indicating that students struggle with solving counting problems. In this paper, the use of the problem-solving strategy of solving smaller, similar problems is tied to students' facility with sets of outcomes. Drawing upon student data from semi-structured interviews in which post-secondary students solved counting problems, evidence is given for how numerical reduction of parameters can allow for a more concrete grasp of outcomes. The case is made that the strategy is particularly useful within combinatorics, and avenues for further research are discussed.

Keywords: Counting, Combinatorics, Discrete Mathematics, Problem Solving

Combinatorial topics have relevant applications in areas such as computer science and probability (e.g., Jones, 2005; Polaki, 2005), and they provide worthwhile contexts in which students can engage in meaningful problem solving. In his undergraduate textbook *Applied Combinatorics*, Tucker (2002) says of his counting chapter, "We discuss counting problems for which no specific theory exists" (p. 169), emphasizing combinatorics as an ideal setting for fostering meaningful problem solving and rich mathematical thinking. Research indicates, however, that students face difficulties with combinatorial concepts, and this is certainly true at the undergraduate level (e.g., Lockwood, 2011a; Eizenberg & Zaslavsky, 2004; Godino et al., 2005; Hadar & Hadass, 1981). Given such difficulty, there is a genuine need for researchers to identify specific areas of struggle for students and to attend to potential ways in which students may improve in their combinatorial problem solving.

In this paper, one particular aspect of combinatorial problem solving is emphasized – the use of the problem solving strategy of solving smaller, similar problems. The research goal addressed in this paper is to highlight the value of this strategy in the domain of counting, particularly because such work can effectively facilitate students' uses of sets of outcomes.

To clarify what is meant by "smaller, similar problems" in this paper, consider the following. In a counting problem, there are typically a number of conditions that specify what the problem is asking. These conditions refer to the rules or limitations that must be met in a given problem. Some of these might be numerical in nature (e.g., the specific number of letters in a password), but others might refer to non-numerical conditions (e.g., the fact that repetition of letters is allowed in a password). For clarity, numerical conditions are referred to as *parameters,* and non-numerical conditions are referred to as *constraints*. In relation to a given problem, any other problem that a student might attempt to solve is called "smaller" if it reduces one or more of the parameters in some way, and is called "similar" if it generally maintains the constraints of the original problem. As an example of the smaller, similar problem strategy studied here, an original problem may be, "How many 6-character license plates can be made, where each character can be any capital letter or any numerical digit? Repetition of characters is allowed." The primary *parameters* of the problem are that there are a certain number of characters (six), and that there are a certain number of options for each character (36; 26 capital letters and 10 numerical digits). *Constraints* include the fact that repetition of those characters is allowed, and, because license plates are being counted, the order of non-identical characters matters (ABCDEF would be a different license plate from FEDCBA). A smaller, similar problem would be "How many 4-character license plates can be made, where each character must be A, B, or C, where repetition is allowed?" because both of the parameters (the number of characters and the number of choices for each character) are numerically reduced, and the overall constraints of the problems are maintained (that repetition is not allowed and that license plates are the objects being counted). A problem such as "How many 6-character license plates can be made, where each character must be A, B, or C, where repetition is allowed, that contain exactly two vowels?" reduces one of the parameters (the number of choices for each character) and changes the constraints of the problem (introducing the additional requirement that the license plate must have exactly two vowels), so it would be considered "smaller" but not "similar." It is noteworthy that the numerical reduction of the parameters may affect certain mathematical properties of those conditions, sometimes in unexpected ways. For instance, if a student reduced a composite number (such as 100) to a smaller prime number (such as 3), there may be some unexpected properties about the small prime number that would not carry over to the original problem.

Literature Review and Theoretical Perspective

When talking about reducing problems to smaller, similar problems, an important issue is whether the similarity exists in the eyes of the student or the researcher. Lockwood (2011b) suggests *actor-oriented transfer* (AOT) as a valuable lens through which to examine students' combinatorial problem solving. Lobato (2003) introduced the notion of AOT, a methodological perspective in which the researcher focuses on student-generated connections between problems, and not on connections that the researcher may expect. In this study, in line with Lobato (2003), student-generated similarities are emphasized, not relationships that the researcher determined to be similar.

The specific strategy of solving smaller, related problems has been alluded to by problem solving researchers like Schoenfeld (1979, 1980), Polya (1945), and Silver (1979, 1981). Polya discusses this strategy in terms of "discovering a *simpler analogous problem*" (p. 38, emphasis in original) and suggests that solving such a problem provides a model to follow when solving the original problem. Schoenfeld (1979) conducted a study examining the effectiveness of explicitly teaching problem solving heuristics, and considering "a smaller problem with fewer variables" (p. 178) was one such heuristic that he examined. Schoenfeld's and Polya's attention to such a strategy suggests that it could be valuable for problem solving across a variety of mathematical domains, but the strategy itself does not seem to have been explicitly targeted as an area of study.

Relatively little has been investigated about the use of smaller problems in the domain of counting problems. Eizenberg and Zaslavsky (2004) allude to such a strategy in their work on undergraduate students' verification strategies on combinatorial tasks. One of those strategies they identified was "Verification by modeling some components of the solution" (p. 26), and one aspect of such verification involved applying "the same solution method by using smaller numbers" (p. 26). Eizenberg and Zaslavsky provide an example of an expert mathematician effectively using this strategy, but they note that while the strategy "could be very powerful…it requires deep structural consideration." They go on to say that "We speculate that although it may seem natural to students to employ this strategy (as indeed some tried to), applying it correctly needs direct and systematic learning" (p. 32*)*. In addition, Maher and her colleagues (e.g., Maher & Martino, 1996; Maher, Powell, & Uptegrove, 2011) have emphasized ways in which carefully chosen tasks may foster sophisticated reasoning in even very young students.

Their work has shown evidence that students generalize using smaller problems, which suggests that meaningful representations and structures can form key mathematical relationships for students. However, researchers have not yet tied the particular strategy of solving smaller, similar problems to developing students' facility with sets of outcomes in a meaningful way.

Additionally, the paper is framed within Lockwood's model of students' combinatorial thinking (Lockwood, in press), which explores the relationships between formulas/expressions, counting processes, and sets of outcomes. Lockwood notes that different counting processes impose different structures on the set of outcomes and suggests that significant progress can be made for students as they gain facility with the set of outcomes. Elsewhere (Lockwood, 2011a), Lockwood advocates for the value of a set-oriented perspective toward counting, suggesting that students may benefit from fundamentally viewing the enumeration of outcomes as an intrinsic aspect of the activity of counting. Focusing on the problem solving strategy as it relates to Lockwood's view on outcomes in particular is the major focus of this paper.

Methods

Twenty-two post-secondary students (12 undergraduates and 10 graduate students) participated in individual, videotaped, 60-90 minute semi-structured interviews in which they solved combinatorial tasks. All of them had taken at least one course in discrete mathematics, and some had taken courses in combinatorics or graph theory. Semi-structured interviews typically involve "an interview guide as opposed to a fully scripted questionnaire" (Willis, 2005, p. 20), and this methodology fostered flexibility that allowed the interviewer to adapt to students' responses. The structure of the interviews was first to give students five combinatorial problems to solve on their own, during which time they were encouraged to think aloud as they worked. After they had completed work on the five problems, the students subsequently returned to a subset of these problems, and they were presented with alternative answers to evaluate. The motivation for this design was based on a desire to put students in a situation in which they had to evaluate incorrect but seemingly reasonable answers. Examples of such incorrect but reasonable answers are discussed below, which would help make the case for the rationale of this methodology. Further details of the study can be found in Lockwood (2011a).

Tasks. The students in the study were given the five tasks listed below.

- **The Passwords problem**: A password consists of eight upper case-letters. How many such 8-letter passwords contain at least three Es?
- **The 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?
- **The Groups of Students problem**: In how many ways can you split a class of 20 into four groups of five?
- **The Test Questions problem**: Suppose an exam consists of 10 questions, and you must answer five questions. In how many ways can you answer five questions if you must answer at least two of the first five questions?
- **The Apples and Oranges problem**: How many different nonempty collections can be formed from five (identical) apples and eight (identical) oranges?

Two problems, the "Passwords" problem and the "Groups of Students" problem, are presented in detail, and any student data subsequently presented comes from work on one of these two problems. Below, both a correct answer and an incorrect answer of each problem are provided; this should facilitate subsequent discussion and give an example of the situation that students faced as they compared alternative answers. The reader may note that in these problems a) the outcomes that are being counted are relatively abstract (that is, they are not concrete objects like books or passwords, but rather are partitions of a class), b) the cardinality of the set of numbers (the answers) are quite large, and c) there are two answers that are incorrect but that could make sense (indeed, each incorrect but reasonable answer below was actually given by students in the interview). Each of these aspects of the problem may make them particularly well suited to facilitate smaller cases.

The Passwords Problem. The Passwords problem states, “A password consists of 8 upper-case letters. How many such 8-letter passwords contain at least three Es?” The at least constraint is noteworthy. Because of this constraint, one solution is to break the problem into cases,[1] in which the passwords contain three, four, five, six, seven, or eight Es. For any of those cases, the number of passwords containing k Es is found by choosing spots for those Es to go (there are $\binom{8}{k}$ way to do this), and then filling in the remaining $8-k$ spots with any of the 25 letters that are not E. Therefore, a correct result (subsequently referred to as Expression PC for Passwords Correct) is

$$\sum_{k=3}^{8}\binom{8}{k}25^{8-k}=\binom{8}{3}\cdot 25^5+\binom{8}{4}\cdot 25^4+\binom{8}{5}\cdot 25^3+\binom{8}{6}\cdot 25^2+\binom{8}{7}\cdot 25^1+\binom{8}{8}\cdot 25^0 .$$

There is also a tempting solution that does not involve a case breakdown, which reflects a subtle error. Namely, the answer $\binom{8}{3}\cdot 26^5$ (subsequently referred to as Expression PI for Passwords Incorrect) can be argued by first choosing where to put three Es (there are $\binom{8}{3}$ ways to do this), there are guaranteed to be at least three Es in the password. Therefore, the remaining five letters could be any letter, including an E (hence, 26^5). However, the problem with this answer is that some particular solutions get counted more than once. For example, the solution $\underline{E}$ $\underline{E}$ $\underline{E}$ $\underline{A}$ $\underline{B}$ $\underline{E}$ $\underline{E}$ $\underline{E}$ gets counted multiple times, both when the first three Es were chosen ($\underline{E}$ $\underline{E}$ $\underline{E}$ _ _ _ _ _) and the rest of the word was filled in with ABEEE, and then again when the last three Es were chosen (_ _ _ _ _ $\underline{E}$ $\underline{E}$ $\underline{E}$) and the rest of the word was filled in with EEEAB. To give a sense of the magnitude of these expressions, Expression PC yields 575,111,451, while Expression PI gives 665,357,056.

The Groups of Students Problem. The Groups of Students problem states, “In how many ways can you split a class of 20 into 4 groups of 5?” A correct answer to this problem (subsequently referred to as Expression GC for Groups Correct) is

[1] In the strategy of breaking a original problem into disjoint sub-problems, the word “case” “case” suggests one such sub-problems. This is a separate use of the word “case” than in the rest of the paper, in which a smaller “case” is meant as a smaller instance of the original problem.

$$\frac{\binom{20}{5}\cdot\binom{15}{5}\cdot\binom{10}{5}\cdot\binom{5}{5}}{4!}.$$

To arrive at this solution, five students can be chosen to be in a group (there are in $\binom{20}{5}$ way to do this), then 5 of the remaining students are chosen to be in another group, $\binom{15}{5}$, then five more to be in a group, $\binom{10}{5}$, and then finally the last five to be in a group, $\binom{5}{5}$. However, the product must be divided by 4 factorial because the groups are not meant to be labeled or distinguished in any way – there is not a Group 1, Group 2, Group 3, and Group 4.

In order to see the need for division by 4 factorial, particular outcomes can be considered. Suppose the first group that was chosen consisted of kids A, B, C, D, E, (call it Group A), the second group, kids F, G, H, I, J (Group F), the third group, kids K, L, M, N, O (Group K), and the fourth group, kids P, Q, R, S, T (Group P). Then this solution could be written as AFKP, and it denotes the order in which the groups were chosen. However, suppose that each of the groups were the same (Group A still consists of A, B, C, D, E, etc.), but that they were chosen in a different order. That is, suppose the solution instead was KAFP or PAKF, representing the same groups that were picked in different orders. The problem only asks for groups of children; there is no first group, second group, etc. Therefore, if the particular division of the class is AFKP, it is the same division as KAFP. In fact, such a division occurs exactly 4 factorial ways (the number of ways to arrange the sequence of letters A, F, K, P). Division by 4 factorial thus ensures that each solution gets counted once, as it should be. A typical incorrect solution (subsequently referred to as Expression GI for Groups Incorrect) is

$$\binom{20}{5}\cdot\binom{15}{5}\cdot\binom{10}{5}\cdot\binom{5}{5},$$

where the division by 4 factorial is neglected. Again, to get a sense of the size of these numbers, Expression GC gives 488,864,376 and Expression GI yields 11,732,745,024.

Data Analysis. Initial data analysis involved transcription of the videotape excerpts. Then, in line with grounded theory (Strauss & Corbin, 1998), the data was carefully coded for phenomena that could be organized into themes. As this was a part of a larger study (details can be found in Lockwood (2011a), and other findings are reported elsewhere in Lockwood (in press)), the entire set of transcripts and videotapes had already been reviewed, and noteworthy recurring ideas had been coded. The notion of smaller cases presented in this study was one such emergent theme, and episodes that represented instances of students' work with smaller cases were flagged. These episodes were watched again, and pertinent phenomena were categorized. What emerged were relevant instances of student work that highlighted students' uses of smaller, similar problems, particularly as these related to students' work with sets of outcomes.

Results

The use of smaller, similar problems arose a total of 15 times. Table 1 below shows students[2] who used the strategy, which problems they worked on as they used the strategy, and whether the problem properly reduced numerical parameters while maintaining constraints. In this table, at most one instance per student per problem is counted – there is a 1 in a cell if the student used smaller cases on a given problem in any capacity. The reader may note that with the exception of Owen's work on the Cards problem, all of the instances of small cases arose in the context of either the Groups of Students problem or the Passwords problem. While students solved three other problems in the interviews, the small cases strategy did not arise in those contexts. One possible explanation for this is that the other problems involved numbers that were relatively small in magnitude. The correct answers to the other problems were 153, 226, 53, compared to 575,111,451 for the Passwords problem and 488,864,376 for the Groups of Students problem.

Student (Undergraduate or Graduate)	Problem	Worked on a smaller, similar problem
Aiden (U)	Groups of Students	1
Anderson (U)	Groups of Students	1
	Passwords	1
Jenny (G)	Passwords	1
Keith (U)	Groups of Students	1
	Passwords	1
Matthew (G)	Groups of Students	1
Mia (G)	Groups of Students	1
Mike (U)	Groups of Students	1
Owen (G)	Cards Passwords	1
	Groups of Students	1
Paige (G)	Groups of Students	1
	Passwords	1
Zach (G)	Groups of Students	1
	Passwords	1
	TOTALS	14

Table 1: Students who used the strategy of solving smaller problems

It is also noteworthy that the strategy was relatively infrequently used. While for some students this seemed to be a strategy that they regularly employed (these students said as much), many of the students did not use it at all. That is, although small cases arose a total of 15 times, only ten (six graduate and four undergraduate students) of the 22 students drew upon the strategy, meaning 12 students never used small cases in their interviews. However, in spite of the fact that the strategy seems to have been underdeveloped and underutilized among the students, the strategy overwhelmingly helped those students who chose to implement it. Given students' overall difficulties on the problems, then, and given the fact that the use of smaller cases seemed to help some students, the strategy is worth examining as a potentially powerful aspect of

---

[2] All participant names are pseudonyms.

combinatorial problem solving. The major result of this paper is that the use of smaller, similar problems gave students greater access to sets of outcomes. That is, students were able more clearly to identify and manipulate outcomes when parameters were reduced, which seemed to give them traction as they solved the problems. Four examples from the data are given below, all of which emphasize how the smaller cases facilitated systematic listing of outcomes, which ultimately shed light on the original problem. The first two examples show how the reduction of parameters facilitated listing of outcomes, which then led to the detection of an overcount and a thus a justification of a correct solution. The other examples similarly show how students arrived at a justification, but with the added feature of emphasizing how students used the outcomes to detect a pattern that they could then check with each given expression.

Facilitating listing of outcomes to detect an overcount and to justify a solution. In her work on the Groups of Students problem, Mia had initially arrived at the incorrect Expression GI. Upon revisiting the problem, she was asked to evaluate and compare her answer with the correct Expression GC. Mia was thus in a position of comparing two different expressions to determine which was correct. Mia had some initial intuition about the role of 4 factorial, but she decided to attempt a smaller case in order to be sure. She worked through a smaller case of dividing six people into two groups of three. She first wrote down A, B, C, D, E, F to represent the people, and she wrote two circles with three dashes each in them. She noted that if she applied her initial method to the smaller case, she would get $\binom{6}{3}\cdot\binom{3}{3}=20$, and she stated that 20 "would not be too bad to write out." Mia then stated that if she applied the other expression to the situation, she would get 10. This is quite noteworthy, as by reducing the parameters Mia changed the numerical comparison from comparing 11.7 billion to 488 million to comparing 20 to 10. This enabled her actually to list outcomes and to determine whether she would get 20 or 10 as her answer for the small case. This is an instance in which Mia, in the context of the smaller situation, computed totals according to both possible expressions and compared the two. The smaller numbers allowed her to begin to write particular examples of outcomes, whereas within the original context this could not have been done feasibly.

Mia then wrote out divisions of six students into groups of two, and she wrote out ABC DEF, then CEF ABD, and then CDF ABE as possible divisions of the students. She paused and then wrote DEF ABC, and something significant happened: Mia noted that this was the same outcome as something she had already written – that is, ABC DEF was the same as DEF ABC (Figure 1 below). It seems that the smaller case (and specifically the smaller numbers) enabled Mia to write out some particular outcomes that she otherwise would not have been able to do (she had not written out such outcomes in her work on the original problem).

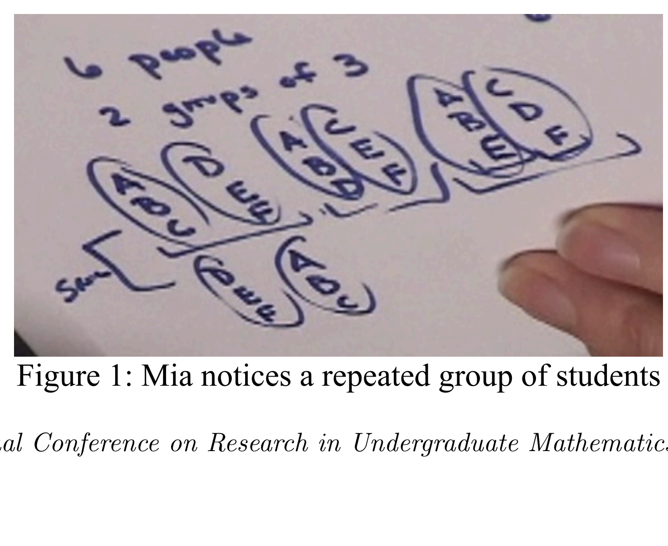


Figure 1: Mia notices a repeated group of students

M: Um, alright A, B, C and then that forces DEF here. So that's one. ABD CEF. ABE and CDF. Hmm. Let's see, oh right, because the first 3 could have been DEF, and then I would have been forced to put ABC in this group, but that's really the same, so these [referring to ABC DEF and DEF ABC] are really the same...Okay, yeah, I think that this double counts because if I just choose 3 people, it could have been A, B, and C, and then that forces DE and F in the second group. But, let's say the first three people were DEF, that forces ABC in the second group, and that's exactly the same, just, it doesn't matter, there's, there's, ABC are in a group and DEF are in a group.

Mia explained the overcounting by referring to her initial solution; she identified two outcomes as being "exactly the same," as her language underlined above indicates. Mia was ultimately able to articulate why division by 4 factorial in the original problem made sense and to identify the correct solution.

In another example, we see how a careful examination of a smaller case can shed light on key differences between two seemingly-similar counting processes. In his work on the Passwords problem, Anderson found himself in a situation of comparing Expression PC and Expression PI, though he did not know which one was correct. He took some time to consider it, and as he talked through the expression, he justified to himself why it might make sense; he explained the (incorrect) counting process and said, "based on how it works" he felt that it could make sense.

As Anderson thought more about why the discrepancy occurred, he decided to truncate the situation to examining the number of 4-letter passwords that contain at least three Es, rather than 8-letter passwords that contain at least three Es. Specifically, this came about as he had written down slots and had put 1s where an E would go, and 25s where a non-E would go. After writing a couple of these slots, he decided to look instead at only four slots instead of dealing with all eight (see Figure 2). He noted that there were 25 options for any placement of three Es, and he explained again that there were $\binom{4}{3}$ ways to place three Es in the four slots, and then 25 choices for the remaining letter. And then he considered the case in which all four letters were Es, and he stated that there was just one way to do that (this gave him an initial answer of 101).

Then, Anderson engaged in some numerical computation for several minutes. He tried the 4-letter problem with the second alternative answer (Expression PI). When he did this, he arrived at $\binom{4}{3}\cdot 26$, which gave an initial answer of 104. Expression PC had given him an answer of 101, and thus he noted, "The difference is already there."

A: Uh, if I was going to translate to the second one then it becomes, uh, in the four case it becomes $\binom{4}{3}$ equals four, but since the last one doesn't matter, it's 26, which is, uh, blah, 104. Well, let's see. Uh, what was the total on this? It was 100, and 1. **The difference is already there.** It's 101, versus 104. Um, okay well since there's already a difference here in the half case, where I'm only taking four letters instead of eight.

This part of the episode was quite significant. Anderson realized that the discrepancy between the expressions existed even in the smaller problem, and ultimately this enabled him to focus on the smaller case. This numerical difference of 101 and 104 was small enough for him to be able to consider in detail, and he proceeded to examine the difference here more closely.

His subsequent work was key for him. First, he said, “I have E, E, E, A through Z. Which is equal to 26.” And he wrote down E E E A-Z. Then, he said, “Then I have another E, E, A through Z, E, which is another 26,” and he wrote E E A-Z E. Then he said, “And since I do this four times, I have 4 times 26, which is 104, okay, which would suggest that the second one (Expression PI) is correct” (see Figure 2 below, the right half of the paper).

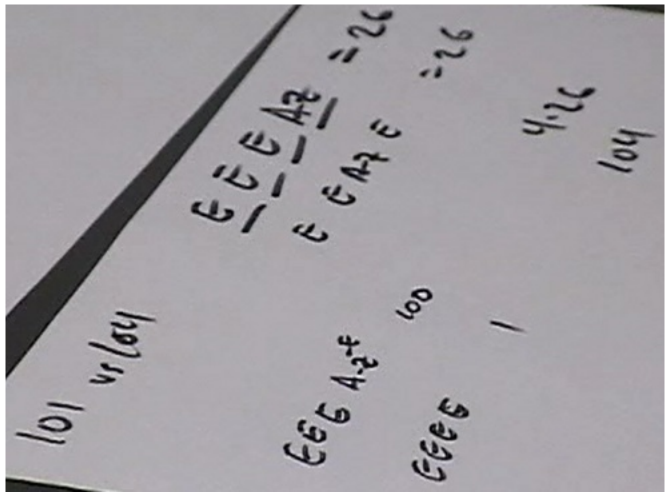


Figure 2: Anderson writes out a smaller version of the Passwords problem

Then, nearby (the left half of the paper in Figure 2), he wrote out E E E A-Z-E, and he said, “I have 3 Es, then I set them to any 25 letters, so let’s see, A through Z minus E. And so I have 100 different ways to do that.” Then he wrote E E E E and noted “But then I also have 4 Es, and there’s only one way to do that.” His reflection on this discovery is seen below.

A: Oh, there we go, that’s where the difference is. So the difference is, um, yes there’s 26 different ways to arrange it so that the first 3 letters are Es, and then the last one can be any of the 26 letters. And then there’s another way to arrange it so that the first 2 and the last letter are Es, and the 3rd letter is any letter between A and Z, **except if the third letter is an E, it’s exactly, it’s the exact same case as if the E was the last letter in the first case, which means it’s counting multiple passwords twice.**

In the excerpt above, Anderson identified a particular password (the all Es password) that was counted too many times by the incorrect solution. After this, in response to follow up questions, Anderson articulated why it took him a while to explain what was wrong with the problem. Simply writing down the range (either A – Z or A – Z excluding E) was not enough for him to see what was happening. His brain was “too lazy to come up with a specific example,” and it was not until he came up with that particular example that he could identify the error. This is an insightful self-reflective statement. He was able to identify that a key step in seeing the issue was writing down a specific example instead of just writing down a range of choices. The fact that he was in a smaller case (with fewer letters and slots to consider) seemed to draw his attention to outcomes, thus enabling him to see the all Es password that would be overcounted.

A: And I was like, oh, the problem – the two methods still come up with different answers, so something must be off on some fundamental level somewhere. Uh, so I realized, well,

> since my brain's not all that math oriented, I guess I'll just like write it out and see where I go, uh, so let's come up with a few examples, so I was like EEE, and I was like, well, my brain's too lazy to come up with a specific example, so I guess I'll just write down the range, and then I should be okay…. Um, and I guess it's that step that my brain kept skipping due to laziness, (chuckles) that made me overlook that one problem.

After this discussion with the smaller case, Anderson was able to use his work in the small case to make sense of why Expression PC was correct in the original problem. His work through the smaller case and his focus on one particular outcome that was overcounted were vital aspects of Anderson successfully evaluating the alternative solution and deciding on the correct answer.

Facilitating pattern generation. Students' work with smaller cases also seemed to facilitate pattern generation. The two students presented below focused on developing patterns as they listed outcomes. This pattern generation and examination was one way in which the students organized and utilized the outcomes of the smaller cases in order to make inferences about the original problem (and, ultimately in order to determine which original expression was correct).

As an example, we consider Paige's work on the Passwords problem. Paige was utilizing a correct total-minus-bad approach, in which she wanted to subtract passwords with exactly zero, one, or two Es from the total number of 8-letter passwords. She quickly determined the number of passwords with exactly zero and exactly one E. As the final part of her total-minus-bad approach, she needed to figure out how to count the number of 8-letter passwords with exactly two Es. Paige realized that in order to solve this part of the problem, she needed to figure out the number of ways to place two Es among eight spots. She felt confident about how to fill in the remaining letters (in 25^6 ways), and so her success on the problem momentarily hinged on one particular part of the problem – placing the two Es in eight spots. Paige suspected that the number of ways to do so involved a permutation, ${}_8P_2$, but she was not sure. (Note, this is not correct. She should use combinations, and not permutations, to place the Es.) In order to check whether this permutation formula was correct (thus allowing her to solve the problem), Paige utilized a smaller case. Paige's decision to utilize a smaller case for this aspect of the problem was evidenced when she said, "I usually make a much smaller case, and try and see if this formula matches up with a smaller case that I can actually physically count."

Paige then attempted a smaller problem of 3-letter passwords, using the letters A, B, and C, and containing at least two As. Over the next several minutes, Paige's work was characterized by writing out 3-letter words and looking for patterns. It is noteworthy, though, that Paige made an error here. She counted out the number of words that had exactly two As, of which there are six (Figure 3 below shows the 6 words and her computation of ${}_3P_2$). While this is correct, it is presumably not what she meant to count. She should have simply counted the number of ways of placing exactly two As in three spots (of which there are 3). This had an impact on her work, because as she computed the result using the formula she was testing, ${}_3P_2$, it also yielded six. This thus (incorrectly) confirmed her permutation formula for a word of length three.

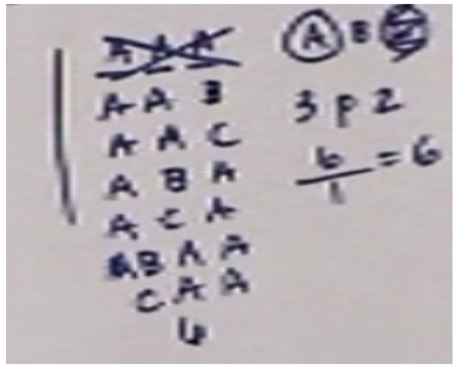


Figure 3: Paige lists 3-letter words with exactly two As

Thinking her permutation formula was correct for the 3-letter case, Paige then proceeded to check for a word of length four, which took considerably more time. This activity involved detailed listing of 4-letter passwords and looking for patterns as she went. In so doing she realized that regardless of where she placed her two As in the 4-letter password, she always had nine options of how to arrange the remaining letters (which could be Bs or Cs).[3] We see her language below, and she made a significant realization that actually she wanted to choose spots for the As.

P: Hmm, how can I make that better? So how many different ways could I place, 'cause those were all 9's. So really I just want to know how many different ways I could place 2 As. Oh, well so, if I think of like the slots as like numbered, maybe,

I: Okay.

P: Um, it's like I want to choose the slots, And the order, hmm, does order matter here, I think it might be that I want to, like, know, like I have 4 slots, and I want to know the number of ways I can choose 2 of those slots to put the As in, but then I think I need to divide that answer by 2, because I can't tell the difference between the As.

This was an important step for her, ultimately allowing her to figure out that the number of passwords with exactly two Es was $\binom{8}{2} \cdot 25^6$ (and not ${}_8P_2 \cdot 25^6$), and thus to arrive at the correct answer in the original problem. Paige's successful utilization of this strategy involved listing outcomes of a smaller problem, and this allowed her to realize that she wanted combinations, and not permutations, for the number of ways to place Es in the password. This work enabled her to generate the correct answer and ultimately be successful on the Passwords problem.

As a final example of pattern generation related to the listing of outcomes, Anderson spent considerable time and energy listing out particular outcomes in the context of smaller cases in

[3] It is not clear why Paige made the transition from incorrectly counting words in the 3-letter case to correctly counting the ways of placing As in the 4-letter case. Perhaps because she could not write the entire words as easily as she could in the 3-letter case, she was more inclined to separate out the placing of the As and the filling out of the rest of the word (the Bs and Cs) as two separate processes. Thus, she was less inclined to inadvertently count 4-letter *words* with two As, but rather she could count the ways of placing two As within a 4-letter word.

the Groups of Students problem. To scale back the original problem, Anderson decreased the number of groups and the number of options to make the problem more tractable. He started with dividing a class of four into two groups, and through listing found that there were three ways to do this. He then increased the problem to a class of six being split into two groups, and through careful systematic listing he found that there were 10 such possibilities. At the heart of this work was pattern recognition – he was searching for a pattern in the numbers in order to generate the correct answer, noting, "I can see where this pattern is going." He made an initial guess at what the general formula might be (see Figure 3, as he wrote, he said, "the number of students choose the size of the groups, divided by the number of groups").

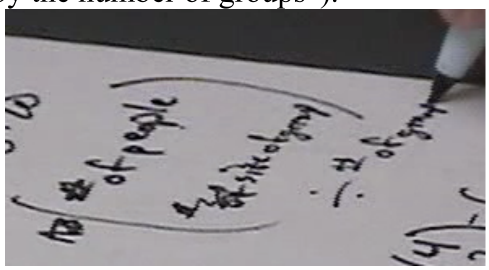


Figure 4: Anderson's initial guess based on pattern generation with outcomes

While this initial formula is incorrect, what is important here is that he recognized that some division needed to take place to prevent overcounting. By having written out so many outcomes in the smaller case, it seems that he was attuned to the fact that overcounting could arise. He says, "Um, in order to answer the original question, though, how many different ways I can split them up, I have to divide it by the number of groups, because in this case, AB CD would be the same thing as CD AB. Sure the groups have switched places, but they're still in separate groups, and the group composition would be the same," which highlights his attention to the outcomes. While he ultimately ran short of time in the interview before Anderson could entirely finish this problem in his initial attempt, he was on a very productive path toward making meaningful progress, particularly in identifying and listing outcomes. Indeed, in his work we see a focus on outcomes. When Anderson revisited the problem and was presented with the two common answers, he related them to the work and patterns he had generated initially, and he made sense of the division by four factorial almost immediately,[4] "Dividing by 4 factorial, I want to say that's probably because there's 4 groups, and factorial helps us cover the different orders of the groups. So like Group A, Group B would be the same as Group B, Group A." When it came to making sense of the solutions, then, it seemed as though his detailed work of systematic listing and looking for patterns was instrumental in helping Anderson understand the problem generally and justifying why the one answer overcounted and the other was correct.

Discussion

In prior work, Lockwood (2011a; 2012) has indicated the importance of considering sets of outcomes for students as they count, suggesting that much benefit could be afforded by explicitly utilizing sets of outcomes in the activity of counting (see also Hadar & Hadass, 1981; Polaki, 2005). The findings in this paper build upon this notion, suggesting that the use of smaller cases enabled students to engage with sets of outcomes through systematic listing. In this section, two aspects of the strategy of solving smaller, similar problems are discussed. First, potential pitfalls

[4] Seven of the 22 students were never able to reconcile the need for the division by 4 factorial.

of the strategy are presented, and second, aspects that make the strategy particularly well suited to counting problems are described.

The above examples of student work highlight various ways in which students used the problem solving strategy of solving smaller, similar problems, and these results emphasize the potential benefit of such a strategy for solving counting problems. However, there were also instances in the data that suggested the strategy could lead to negative consequences if it was not implemented properly. There are three potential pitfalls that arose, and these are presented to point out what could happen if the strategy is not carried out with care and precision.

First, unexpected mathematical properties can arise between the smaller problem and the original problem. As students numerically reduce a parameter in order to create a smaller problem, some typically unforeseen mathematical property can arise (or disappear) that had been absent (or present) in the original problem. This issue arose in Paige's work, when she evaluated the correct alternative answer to the Groups of Students problem. Expression GC has a 4 factorial in the denominator, while Expression GI does not, and when Paige compared the two answers, she sought to examine whether division by 4 factorial, or simply division by 4, was correct. Paige had reduced the problem to splitting into two groups; had she applied Expression GC to such a case, it would have meant dividing by 2 factorial. However, two and 2 factorial are equal, and simply by looking at one small case she could not clearly tell whether she was dividing by two or by 2 factorial (and thus whether she should divide by four or 4 factorial in the original problem). While the reduction of the problem from dealing with four groups to two groups undeniably made the outcomes easier to list, a specific property about the number 2 (that does not apply to the number 4) affected Paige's ability to generalize from the smaller case. Fortunately, Paige was able to recognize 2's special property, and she sought other means of reasoning through why 4 factorial made sense. This example highlights the fact that some mathematical objects have properties that may cause them to behave in certain ways, and reducing parameters can enhance some properties or inhibit others. It takes mathematical knowledge and experience to foresee how certain numerical properties could affect the constraints of the problem, and this is issue is something of which students ought to be aware.

Second, the relationship between the smaller problem and the original problem can be lost. Students should not become so involved in a smaller case that they are unable to relate it back to the original problem. For example, in her work on the Passwords problem, Jenny spent a considerable amount of time working through a small example to make sense of a particular part of the problem. At one point, she was trying to remind herself of the ways of placing the remaining non-E letters, and to address this issue she looked at a case of only 3 letters. In the course of this work, she drew a very detailed tree diagram and spent a considerable amount of time working through that diagram. Eventually, this work had the effect of distracting her from what she had initially been doing. She had difficulty recalling how her endeavors particularly related to what she was trying to accomplish. This is not to say that detailed and lengthy forays into smaller cases are always bad – Anderson, for example, showed instances of very effective work with small cases that could be called tangential. However, such work is only effective insofar as it can be directly related back to the problem at hand.

Third, the smaller problem can be worked carelessly, without precision and attention to detail. This is not new advice in the problem solving literature; Polya emphasizes this in his famous book, How to Solve It (e.g., pp. 68-69). This is particularly noteworthy in counting problems, though, in which numbers can be large, and maintaining the integrity of the original problem can be difficult. In the above examples of successful implementation of small cases, these students' work was marked by very precise, detailed work. The students invested time in figuring out the problem, and they were willing to try multiple cases and to engage in systematic listing. While these are certainly desirable qualities for mathematics students in a variety of settings, these are particularly important attributes when dealing with counting problems.

Despite such potential pitfalls, the use of this strategy seems to be especially beneficial in the context of counting problems. In particular, the use of a smaller problem allowed for work with the set of outcomes that might not otherwise have been attainable. The nature of counting problems makes them particularly appropriate for the strategy of using smaller, similar problems. Specifically, in counting problems, sets of outcomes are often so large that they can be difficult to conceive of and manage. Smaller cases can reduce the magnitude of such sets and can make the problems and the solution sets more accessible.

Additionally, it is very important for a student to be able to articulate what he or she is trying to count, and smaller cases can facilitate such activity. In some counting problems the objects being counted can be quite difficult to articulate (in the Groups of Students problem, an outcome is one partition of 20 students into four groups of five; this necessitates coordinating a number of factors – 20 distinct students, what such a division might look like, and how they might be divided to create a desirable division). Smaller cases are particularly useful because they allow not only for the magnitude of the outcomes to be reduced, but often also for the outcomes themselves to be easier to identify. However, students must be aware of the fact that reducing a problem can introduce unexpected mathematical properties, and care must be taken when manipulating the original problem. Finally, solutions to counting problems are notoriously difficult to verify (Eizenberg & Zaslavsky, 2004), particularly because errors in seemingly reasonable counting problems can arise in subtle ways. However, as shown in the examples above, solving smaller, similar problems can facilitate work with outcomes, which may serve as a powerful means of verification.

Conclusions and Avenues for Further Research

The results discussed above highlight the fact that solving smaller, simpler problems can allow students to work with sets of outcomes in meaningful ways. Overwhelmingly, the strategy helped the students who chose to implement it. Given student difficulties with counting problems, the use of smaller cases seems to be a promising strategy that could be useful for students, particularly in this domain. One potential avenue for further research is to relate students' uses of smaller problems with their notions of what determines similarity among problems. That is, by identifying smaller, similar problems, students can be thought of generating particular examples or instances of a given problem type. With increased attention on the role of examples in mathematics education literature (e.g., Bills & Watson, 2008), such an investigation could shed light on how students view examples of particular problem types. Another avenue to pursue is how to generate this problem solving strategy effectively, both among students and among pre-service and in-service teachers. This must be done with care, though, and both students and teachers must be made aware of mathematical complications that can arise when reducing problems.

References

Auerbach, C. & Silverstein, L. B. (2003). *Qualitative data: An introduction to coding and analysis.* New York: New York University Press.

Bills, L., & Watson, A. (2008). Editorial introduction. *Educational Studies in Mathematics*, *69*, 77-79.
Eizenberg, M. M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning*, *6*(1), 15-36.
Godino, J., Batanero, C., & Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60*, 3-36.
Hadar, N., & Hadass, R. (1981). The road to solve combinatorial problems is strewn with pitfalls. *Educational Studies in Mathematics*, *12*, 435-443.
Jones, G. (Ed.). (2005). *Exploring probability in school: Challenges for teaching and learning*. New York: Springer.
Lockwood, E. (2011a). *Student approaches to combinatorial enumeration: The role of set-oriented thinking.* Unpublished doctoral dissertation, Portland State University – Portland, OR.
Lockwood, E. (2011b). Student connections among counting problems: An exploration using actor-oriented transfer. *Educational Studies in Mathematics, 78*(3). 307-322. Doi: 10.1007/s10649-011-9320-7.
Lockwood, E. (2012). Counting using sets of outcomes. *Mathematics Teaching in the Middle School*, *18*(3). 132-135.
Lockwood, E. (in press). A Model of Students' Combinatorial Thinking. To appear in *Journal of Mathematical Behavior.*
Maher, C. A., & Martino, A. (1996). Young children invent methods of proof: The gang of four. In P. Nesher, L. Steffe, P. Cobb, B. Greer & J. Goldin (Eds.), *Theories of mathematical learning* (pp. 431-447). Manwah, NJ: Lawrence Erlbaum Associates.
Maher, C. A., & Speiser, R. (2002). How far can you go with block towers? *The 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.
Polaki, M. V. (2005). Dealing with compound events. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 191-214). New York: Springer.
Polya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.
Schoenfeld, A. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education, 10*(3), 173-187.
Schoenfeld, A. (1980). Teaching problem-solving skills. *American Mathematical Monthly, 87*, 794-805.
Silver, E. (1979). Student perceptions of relatedness among mathematical verbal problems. *Journal for Research in Mathematics Education, 10*(3), 195-210.
Silver, E. (1981). Recall of mathematical problem information: Solving related problems. *Journal for Research in Mathematics Education, 12*(1), 54-64.
Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.).* Thousand Oaks, CA: Sage Publications, Inc.
Tucker, A. (2002). *Applied combinatorics (4th ed.).* New York: John Wiley & Sons.
Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design.* Thousand Oaks, CA: Sage Publications, Inc.

A MULTIDIMENSIONAL ANALYSIS OF INSTRUCTOR QUESTIONS IN ADVANCED MATHEMATICS LECTURES

Melissa Mills
Oklahoma State University

This study is an investigation of the questions that are asked by four faculty members who were teaching advanced mathematics. Observations of each classroom were conducted, and the questions asked by the instructor were analyzed along two dimensions: the expected response type of the question and the Bloom's Taxonomy level.

Key words: Teaching practices, questions, cognitive engagement, Bloom's Taxonomy, lectures, proof courses

1.Introduction

Instructional practices at the undergraduate level have been largely unexamined (Speer, Smith, & Horvath, 2010) and many of the studies that do focus on teaching practice have occurred in lower division courses like calculus, where students are expected to be able to do computations and applications (Thompson, et.al. 2007, Epstein, 2007; Bressoud, 2011). In advanced mathematics courses, the content shifts to formal mathematics, and undergraduates are expected to be able to comprehend and write mathematical proofs. Because the content and expectation of the students is quite different, the teaching of mathematics at this level may need to be examined separately. Previous studies of teaching practice at this level have focused on proof presentation in class (Fukawa-Connelly 2012a; Fukawa-Connelly 2012b; Mills, 2012; Weber, 2004). This study will add to the existing literature on the teaching of advanced mathematics by examining the interactions between the instructor and students.

The participants of this study are four instructors teaching different upper-division proof courses at the undergraduate level. All of these instructors taught using some variation of lecture, meaning that the instructor was primarily standing at the board presenting the material while the students were sitting in desks. Studies that examine teaching often focus on where the instructor's presentation method lies on the continuum of lecture to reform (Sawada, Piburn, Judson, Turley, Falconer, Binford & Bloom, 2002; Steussey, 2006; McClain & Cobb, 2001), but this emphasis on the presentation style tends to gloss over subtle features of teaching practice, namely, differences that may occur within a lecture format. This study will provide a multi-dimensional analysis of the questions that are used by intructors while lecturing in advanced mathematics courses.

2. Research Questions

- How often do instructors who are teaching advanced mathematics using lecture methods interact with their students by asking questions?
- What types of questions are asked by instructors who are teaching advanced mathematics using lecture methods, and what types of responses are expected of students?

3. Literature Review

Lecture is still widely used in undergraduate mathematics instruction, and a majority of instructors believe that lectures can be effective (Bressoud, 2011). However, few studies investigate in detail the teaching practices of instructors using primarily lecture methods. As Krantz (1999) points out, a masterful lecturer may include many different pedagogical moves to connect to his or her audience. Instructors can use examples, give summaries, check for student understanding, or make connections between different topics (McKeachie & Svinicki, 2006). Lecture can also be interactive, incorporating lots of questions that guide students through the material (Bagnato, 1973). In short, there can be significant variation among lecturers.

There are studies that investigate instructor questioning in the classroom, however, most of these studies occur at the K-12 level. In grade-school traditional mathematics classes, a typical questioning pattern involves the teacher *initiating* with a question, the student *responding,* and the teacher *evaluating* the response and proceeding to the next question (Mehan, 1979). These IRE questioning sequences are designed to evaluate whether or not the student knows the answer and often do not require the student to be involved in any mathematical thinking to participate. In fact, there are many who contend that children in traditional mathematics classes can participate in class discussions by merely following the teacher's linguistic and contextual cues (Voigt, 1985, 1989). In these contexts, teacher questions serve to accentuate the power imbalance between the teacher and student.

VanZee & Minstrell (1997) analyze the questioning of an experienced and respected teacher, Minstrell, as he teaches a high school physics class. Minstrell's questions seemed serve a different purpose. He asked questions that focused on what the students were thinking, and he often did not have a predetermined answer in mind. In particular, they noticed his use of "reflective tosses," which are questions that elicit further thinking on the part of the student. The instructor "catches" the students meaning, and "tosses" the responsibility for thinking back to the student. These questions prompt the students to elaborate and reflect upon their thinking and to consider new aspects of the problem at hand.

Teachers of school-level mathematics sometimes use a *funneling* pattern in their questioning. This is when a teacher asks a sequence of narrowing questions to "produce a pre-determined solution procedure preferred by the teacher" in which the "student needs only to generate superficial procedures rather than meaningful mathematical strategies in order to participate." (Wood, 1994, p. 155). Wood (1994) identifies another type of questioning sequence which he calls *focusing.* This is when the teacher asks guiding questions to focus the students' chosen construction path. The teacher's questions serve to turn the discussion back to the student, leaving him/her with the responsibility for resolving the solution.

One study of teaching at the university level investigates how an abstract algebra instructor, Dr. Tripp, used questions to devolve responsibility to students when presenting proofs (Fukawa-Connelly, 2012a). This case study showed that Dr. Tripp used rhetorical questions to model the mathematical thinking involved in creating the proof's structure, and she also asked a large number of questions that solicited student feedback. She often began with a higher-level question but then asked several successive questions in a row until the final question required merely a factual response or re-statement of something she previously said. Though she did devolve some responsibility for proof writing to students, the majority of students' answers stated the next part of the proof or the next algebraic step.

This study will examine the questioning used by four different instructors in advanced mathematics lectures at the university level. The literature on interactions in mathematics classrooms seems to fall into two camps: those that look for overarching patterns in interactions (Fraivillig, 1999; Fukawa-Connelly, 2012a; Henningsen & Stein, 1997; Lobato et al, 2005; Tobin, 1986; VanZee & Minstrell, 1997), and those that classify the types of questions used by instructors and students (Gall, 1970; Sahin & Kulm, 2008; Wood, 1994; Wood, 1999). This study will combine these approaches by analyzing questions on multiple dimensions.

Several existing taxonomies, such as Bloom's Taxonomy, record the cognitive level that the student will need to use to answer the question (Anderson & Krathwohl, 2001; Gall, 1970; Tallman & Carlson, 2012). Gall (1970) points out that there are several types of questions do not fit well into these taxonomies. In particular, in this study, rhetorical questions and general questions that check for student understanding often do not fit well into these taxonomies. Other classification schemes classify questions by the expected products, such as whether the questions require the student to make a choice, give factual information, give reasons for their thinking, or justify their thinking (Wood, 1999; Mehan, 1979). Although this is a reasonable way to catalog question types, the expected response type does not necessarily capture the cognitive processes required to answer the question. This study will consider the expected response types and cognitive engagement as separate dimensions, which will allow for a more in-depth classification of the question types.

In his discussion of Dr. Tripp's questioning patterns, Fukawa-Connelly (2012a) describes some of her questions as "high level" and others as "factual," but this study contributes by providing more specific and detailed descriptions of the questions that were asked in lectures. Bloom's taxonomy has been used to investigate the types of questions that appear on undergraduate Calculus exams (Tallman & Carlson, 2012), but this study differs because the nature of the abstract mathematics in these courses is quite different from the content in Calculus courses. This study also investigates instructors' questions that are used in the context of teaching, not questions that have been constructed specifically for use on examinations.

4. Methods

Video observation data were collected periodically during the regular semester in each of four courses: Geometry, Number Theory, Introduction to Modern Algebra, and Introduction to Advanced Calculus. Six observations of each classroom were conducted, with the camera focusing on the instructor and the chalkboard. The observation days were chosen so that they occurred on instruction days, and were spread out so that they occurred approximately every two weeks throughout the semester.

Semi-structured interviews with each instructor were conducted before the start of the semester, and a follow-up interview was conducted one year after the data were collected, to serve as a member check. The initial interview asked the instructors to describe the different strategies that they would use to help students to understand the proofs that they present in class. The follow-up interview asked some broad questions about how they interacted with their students, but did not directly address the analysis of their questions that is presented in this paper.

The four instructors were given the pseudonyms according to the course that they were teaching. Dr. A taught the algebra course, Dr. G taught geometry, Dr. N taught number theory, and Dr. C taught advanced calculus. All of the instructors were tenured faculty members and experienced teachers. Since the instructors are teaching in different content areas and the content areas or the makeup of the individual classes may have an effect on the instructors' use of questions, they will be considered separate but interrelated case studies.

All of the questions that were posed by the instructor in the observation data were transcribed. Like Van Zee and Minstrell (1997), this included all utterances that had the grammatical form of a question and some alternative forms. For example, when presenting a proof, one particular instructor would often begin a statement and then pause and look at the students. In these situations the instructor appeared to be expecting the students to complete the statement, and so this situation was coded as a question. I also included questions that did not involve the explicit seeking of information, such as rhetorical questions or questions to check student understanding.

For each question, I noted whether or not there was an audible response from the students. Since the video was directed at the instructor and the board and not at the students, I could not identify or code any nonverbal responses to the questions.

Then each instructor question was coded along two dimensions, the Expected Response Dimension and the Cognitive Process Dimension. First, I coded two videos from each instructor according to the descriptions of the categories given in Mehan (1979) and Anderson & Krathwohl (2001). Then, I looked at each category and noted the types of questions that occurred in each category, which helped to solidify the categories and relate the categories to this context. I then created a document that summarized the types of questions that occurred in each category of Bloom's Taxonomy, which can be found in Appendix A. I used this document to code the questions in the remaining videos.

4.1 The Expected Response Dimension

The Expected Response dimension records the type of response that the instructor seems to expect from the students. This includes Mehan's (1979) four types of questions: *choice, product, process, and meta-process*. However, Mehan's categories only classify questions in which the instructor expects a response from the students. The expected response dimension also includes *rhetorical questions* which the instructor does not intend for the students to answer, as well as *comprehension questions*, which can often be answered by the students with non-verbal responses such as making eye contact, facial expressions, or nodding their heads.

4.1.1 Rhetorical Questions. The question was coded as rhetorical if the instructor either answered it himself immediately after posing it, or if he did not wait for the students to respond to the question (and it was not linked to another question).

4.1.2 Comprehension Questions. Questions such as "Does that make sense?" "Ok?" and "Any questions?" were coded as comprehension questions. These often had no verbal response from the students, but the instructor would look at the students and gauge their understanding based on non-verbal cues.

4.1.3 Choice Questions. Yes/No questions or questions where the instructor asked students to choose between two or more options were coded as choice questions.

4.1.4 Product Questions. Questions that require a factual response or short answer were coded as product questions.

4.1.5 Process Questions. When the question required more than just a short answer, or required the student to make interpretations, describe computations, or explain the mathematics content, they were coded as process questions.

4.1.6 Meta-Process Questions. Questions requiring the student to reflect on their thinking or formulate the grounds for their reasoning were coded as meta-process questions. Questions requesting a student to expound upon their response, such as Minstrell's (1997) *reflective tosses,* are included as meta-process questions.

4.2 The Cognitive Process Dimension

Each question was also analyzed to determine its cognitive level based on Bloom's Taxonomy (Anderson & Krathwohl, 2001). Because the Bloom's Taxonomy codes are not specifically tailored to mathematics content, one slight revision to the categories was made. The "remember" category was parsed into "remember" and "apply a procedure" as in Tallman & Carlson's framework (2012). This was done because often in mathematics students can apply a procedure without understanding the mathematics, and therefore the cognitive action of applying that procedure is more at the level of "remembering" than it is "applying." The following are brief descriptions of each level of the taxonomy as found in Anderson & Krathwohl (2001) and Tallman & Carlson (2012). Examples of each type from my data will be presented in the results section.

4.2.1 Remember. Students are prompted to retrieve knowledge from long term memory.

4.2.2 Apply a Procedure. Students must recognize and apply a procedure.

4.2.3 Understand. Students are prompted to make interpretations, provide explanations, make comparisons, or make inferences that require understanding of a mathematics concept.

4.2.4 Apply Understanding. Students must recognize when to use a concept when responding to a question or when working a problem.

4.2.5 Analyze. Students are prompted to break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose.

4.2.6 Evaluate. Students are prompted to make judgments based on criteria and standards.

4.2.7 Create. Students are prompted to reorganize elements into a new pattern or structure.

5. Results

5.1 Frequency of Instructor Questions

The first research question addresses the frequency with which instructors ask questions in advanced mathematics lectures. The average number of questions per minute for the four instructors ranged from 0.69 questions per minute to 1.81 questions per minute. So, it appears that instructors at this level do frequently interact with their students by asking questions. This shows that lectures are not necessarily monologues by the instructor, but do involve interaction with the students.

Table 1: *Frequency of Instructor Questions*

	Algebra	Geometry	Num Thry	Adv Calc
Instructor	Dr. A	Dr. G	Dr. N	Dr. C
# Students	24	9	14	9
# Instruction Days	44	44	44	29*
# Observations	7	6	6	7
Total time observed (min)	323	250	264	536
Question Rate (per min)	0.69	1.05	0.88	1.81
Response Rate	23%	46%	52%	47%

* Dr. C's class meetings were 75 minutes long, while the other three courses met for 50 minutes per class.

The rate of student responses ranged from 23% to 52%, but three of the four instructors had response rates close to 50%. One reason that the response rates may be lower than expected is that this analysis considers rhetorical questions and comprehension questions in the question count, and those types of questions typically do not require an audible response on the part of the student.

Another reason that the response rate may be lower than expected is that the instructors frequently linked questions together, either by re-phrasing the question or asking a related question that reduced the cognitive load, as in Wood's (1994) description of *funneling* or *focusing* questioning patterns. It was often the case that the instructor would ask two or three linked questions before a student would give an answer. For example, Dr. C asked the following questions about the limit of a given sequence: "So, what is the limit in this case? *(pause)* Do you guys remember how to find that limit? *(pause)* When n is very large, what is happening to s_n? *(pause)* Like, when n is 100, what is s_n?" A student answered the final question, which led Dr. C into another line of questioning focused around plotting the terms of the sequence to see if they could determine the limit.

5.2 Examples of Question Type from the Data

5.2.1 Rhetorical Questions. Many of the rhetorical questions that occurred within my data were instances where the instructor was asking a question to either motivate content or to model the thinking that the instructor expects the students to do independently. Some examples include:

- Can you see that the answer is "no?" It does not require the function to be constant.
- So, this implies that cb=da, and what's this implication from? It's symmetry of equality.
- Now, what's your goal? Your goal is to take the set S, which is a subset of the set T, and show that it is finite.

The use of rhetorical questions may be a characteristic of the individual instructor, because the percentage of rhetorical questions varied wildly, from Dr. A using 40.6% rhetorical to Dr. C using 7.1% rhetorical.

5.2.2 Comprehension Questions. This type of question was used by instructors to check for student understanding, and often did not require an audible response on the part of the student. Typical examples of this type of question include:

- Does that make sense?
- Anything seem tricky, or was everything clear?
- Does this make sense, or do you guys have a question?
- Ok?

All of the instructors used comprehension questions in varying degrees, from Dr. G with 3.1% comprehension questions to Dr. C with 16.6% comprehension questions.

5.2.3 Choice Questions. Choice questions are either yes/no questions or questions that require the student to choose between two or more options. Here are some examples of choice questions in different Bloom's levels.

- Recall. "Now, you remember that we had a definition of even and a definition of odd that we had last time, right?"
- Understand. "So, therefore, is this set open or closed by our definition?"
- Apply Understanding. "Would that point be in the boundary of this set S or not?"
- Analyze. "Now, if I take the intersection of open sets, is it open?"
- Evaluate. "Now, does that seem like an easier proof to you?"

Choice questions were used by all of the instructors, and hit all of the Bloom's levels except for *create*, because the nature of a choice question does not allow for the student to create a new structure.

5.2.4 Product Questions. Product questions require a student to give a factual response, or to give the name of a particular mathematical object. Some examples of product questions at different Bloom's levels are:

- Recall. "What property of a graph does (d) describe? There's a word for this…"
- Apply a Procedure. "So, what do I get when I square that out?"
- Understand. "But, just look at the other end of the spectrum. The identity is the littlest subgroup of G, what's the biggest?"

- Apply Understanding. “Well, when you’re writing proofs, usually proofs need to refer back to a definition. What definition am I wanting you to refer to?”
- Analyze. “Can you make a true version of the statement?”
- Evaluate. “I’m sorry, I wrote something wrong. What’s my error right there?”
- Create. “So, now you need to help me figure out the collection of sets.”

All of the instructors used product questions frequently.

5.2.5 Process Questions. Process questions require the student to give their interpretations or opinions, or to describe a mathematical computation, or to explain mathematics content. Some examples of process questions include:

- Recall. “So, the best way to prove uniqueness is to… I mean, what’s the standard way that you prove that there’s only one formula?”
- Apply a Procedure. “And, and, um, you could simplify this a little bit, how would I simplify this?”
- Understand. “We’re not ready to substitute yet until we get it in an external form that models 11.9d that we’re trying to use. So, what algebraic manipulation needs to be done?”
- Apply Understanding. “I have to know something about this, so what would you do? Something you know about Fibonacci numbers.”
- Analyze. “You stop when the quotient is finally zero. But, why are you forced to get a quotient that is zero? You have to detect a pattern in order to…”
- Evaluate. “He proposed to just discard Euclid 5 and replace it with this obviously simpler axiom, "There is a rectangle." What did he accomplish?”
- Create. “Now, what you have to give me, is you have to give me the order in which you are going to list them. What’s the first one?”

All of the instructors used process questions frequently.

5.2.6 Meta-Process Questions. Questions that require a student to reflect upon their thinking or to formulate the grounds for their reasoning are meta-process questions. Some examples are:

- Understand. “So, do, are any of these pictures helpful to you for any particular reason?”
- Apply Understanding. “How did you know that? You didn't work it out mod 7 or something?”
- Analyze. “Absolute value of k. Keep it inside absolute values, don’t get rid of them. Correct. Why did you do that, [student]?”
- Evaluate. “Why would I do that purely by adding and subtracting from both sides instead of adding through by a negative one?”
- Create. “So, any guesses about what I'm about to tell you I want a limit point to be, if that's the notion of limit that I've got in my head right now?”

Most of the instructors did not use meta-process questions frequently. Dr. C used them the most, with meta-process questions making up 3.7% of his total questions. It may be that the main difference between lecture-based instruction and inquiry-based instruction is the frequency of meta-process questions, as hinted by Wood (1994).

5.3 Case Studies of Instructor Questions

Since each instructor was teaching in a different content area with a different group of students, we must consider each classroom as a separate case study. The next sections will give specific information about each instructors’ use of questions, including some excerpts from their interviews to shed light on their perspective in regard to interacting with their students using questions.

5.3.1 Dr. A's Abstract Algebra Class: At this university, the Introduction to Abstract Algebra course is the first proof course that math and secondary math education majors take. Therefore, this course has a wider variety of students, and the students are not as adept in their proof-writing abilities as students in subsequent courses.

In the observations of Dr. A, it was found that 40.6% of the questions that he asked were rhetorical. When asked about this, Dr. A said, "I didn't want to take… you see, that's the thing. I'm impatient. Very impatient. And so, you're absolutely right. I do question, but I'm not particularly wanting them to answer." He also said "I feel uncomfortable… I mean, if I'm asking you a question and you're sitting there and struggling with it, I don't want to just keep putting you on the spot. I want to move on. I'm very uncomfortable embarrassing anybody or making them feel uncomfortable."

Another thing that may have affected Dr. A's questioning was his view of the students in that course. In his follow-up interview, Dr. A said "Well, you know, I thought they [this class] were particularly weak." He also expressed that he didn't think the students were trying very hard, and mentioned that "they didn't do the extra credit problems… which I found very strange." It may be that the perception that Dr. A had of this particular class influenced the types of questions that he asked of the students.

Table 2: *Dr. A's Questions*

	Rec.	App Proc.	Und.	Apply Und.	Anal.	Eval.	Cre.	Total
Choice	1.6%		5.3%	2.0%	2.9%	1.6%		13.5%
Product	2.0%	2.5%	10.7%	1.2%	2.0%		0.4%	18.9%
Process	0.4%		4.1%	1.2%	6.1%			11.9%
Meta-Process								0%
Total	4%	2.5%	20.1%	4.4%	11%	1.6%	0.4%	

Dr. A asked 40.6% rhetorical, and 7.0% comprehension questions. Thus, only 52.4% of his questions were questions that expected a response from the students. His overall response rate was 23%.

5.3.2 Dr. G's Geometry Class: The proof-based Geometry course provides an axiomatic approach to several different geometries, focusing on non-Euclidian geometries. The students in the class were all math or math education majors in their junior or senior years.

Table 3: *Dr. G's Questions*

	Rec.	App Proc.	Und.	Apply Und.	Anal.	Eval.	Cre.	Total
Choice	1.1%		3.4%	0.4%	0.8%	0.4%		6.1%
Product	1.1%		9.5%	9.5%	1.5%			21.8%
Process	0.4%		8.0%	9.2%	14.1%	3.1%	0.8%	35.5%
Meta-Process			0.4%					0.4%
Total	2.6%	0%	21.3%	19.1%	16.4%	3.5%	0.8%	

In the initial interview, Dr. G described the types of questions that he uses when teaching this course. He said "I ask little questions like: 'why is this angle congruent to that angle?' 'What's the next step? What do I do here?' Mostly I ask questions of students and insist on an answer. The worst thing, I think, from a teacher's point of view, is if you try to ask a question, and the class just stares at you, and you can't get any response. That's when you want to throw up your hands and quit."

Most of the proofs that were presented in the Geometry class used proof by contradiction. Dr. G would often go through the proof, asking questions along the way, and then stop and ask the students to spot the contradiction. This type of questions was coded as a Process/Analysis question, and was the most frequent type of question that Dr. G asked.

Dr. G asked 33.2% rhetorical and 3.1% comprehension questions. He had a response rate of 46% overall.

5.3.3 Dr. N's Number Theory Class: The Number Theory course is a senior-level course that is made up of mathematics and math education majors, and covers modular arithmetic, quadratic residues, and other basic number theory topics.

In his interview, Dr. N praised the work of Polya, and talked about how he thought that using questions and interacting with students was the best way to teach. He also mentioned that because of class sizes and time constraints, he was not always able to teach in the way that he thought was best. He said, "well, my uh, usual pattern, well, it slows the class down, so I'm not always afforded the liberty of doing it the way that I like, my usual pattern is to ask questions of different pieces. Polya gives an example in his book of leading a student through a, uh, solving a problem by asking questions. So, that was his belief, that you have to ask questions and have to learn how to ask questions…"

He asked a lot of linked questions, and when students answered incorrectly, he would take some fragment of the students' comment that was correct and re-work it into another question. When asked about that technique, he said, "when you read "How to solve it" by Polya, he has a sample dialogue with a student where he does exactly that. You're not supposed to, you're not supposed to, like, give up, you're supposed to re-word the question, or ask the students to think about a different way of asking about it, or something like that. So, that's intentional."

Table 4: *Dr. N's Questions*

	Rec.	App Proc.	Und.	Apply Und.	Anal.	Eval.	Cre.	Total
Choice	0.9%	0.4%	3.0%		1.3%			5.5%
Product	3.4%	3.4%	6.4%	13.6%	5.1%	1.3%		33.2%
Process	1.7%	0.9%	4.7%	6.4%	14.0%	3.0%	1.7%	32.3%
Meta-Process				1.3%		0.4%		1.7%
Total	6%	4.7%	14.1%	21.3%	20.4%	4.7%	1.7%	

Dr. N asked 10.6% rhetorical and 16.6% comprehension questions in the observation data. His overall response rate was 52%.

5.3.4 Dr. C's Advanced Calculus Class: Introduction to Advanced Calculus is a senior level course that covers the theoretical background for Calculus, such as compactness and the formal definition of limit. The class enrollment consisted of mostly seniors who were math or math education majors.

Dr. C commented that when he lectures he tries to talk to his students and get them to talk back. He said, "I ask them to help me. I think I always ask them to help me, and I think that I have, well, I don't know. Hopefully I always ask them to help me, it depends on how much time we have. But, I want them, one of the things that I want them to do is to do some meta thinking as well as some detail thinking."

When questioning, students would often answer incorrectly, and Dr. C would either re-direct by asking another question, or just remain silent and continue to look at the students. When asked about this strategy in the follow-up interview, he said, "I just try to let them have a chance to think about it. And, I think, if you say to somebody, 'No, you're wrong!' Then their brain doesn't have a chance to process on it. But, if you say, 'Hmmmm…' Then, all the sudden they are still thinking. Well, what's better? Is it better for me to talk or for them to think? It's better for them to think!"

It is also interesting to note that Dr. C asked some meta-process questions. He tended to have a one-on-one dialog with individual students, and in these situations he would sometimes ask them to explain their thinking.

Table 5: *Dr. C's Questions*

	Rec.	App Proc.	Und.	Apply Und.	Anal.	Eval.	Cre.	Total
Choice	2.0%	0.2%	6.3%	2.8%	1.9%	0.3%		14.2%
Product	5.9%	2.5%	13.1%	8.9%	3.7%	1.1%	1.7%	35.9%
Process	0.6%	0.6%	5.3%	6.8%	8.8%	0.2%	1.4%	24.9%
Meta-Process			1.3%	0.3%	1.0%	0.7%	0.3%	3.7%
Total	8.5%	3.3%	26%	18.8%	15.4%	2.3%	3.4%	

Dr. C asked 7.1% rhetorical and 14.1% comprehension questions, with an overall response rate of 47%.

6. Discussion

A main result of this study is that instructors who are using lecture in mathematics courses at this level do interact with their students quite frequently. Also, there is some variation among the participants in the ways that they use questions in lectures. Some instructors use a lot of rhetorical questions to motivate content or to model the mathematical thinking involved in their presentation. Another finding was that these instructors do not use meta-process questions very frequently in their lectures.

This study contributes to the existing literature by giving a quantitative description of instructor questions in lectures at the advanced mathematics level. It also contributes by giving specific examples of questions at different Bloom's taxonomy levels for advanced mathematics. Knowledge about how to use Bloom's Taxonomy to design questions for advanced mathematics lectures could be useful for anyone who is teaching such courses, or for professional development programs.

In Tallman & Carlson's (2012) analysis of Calculus 1 final examinations using Bloom's Taxonomy revealed that almost all of the questions could be categorized in the first three levels of the Taxonomy: Remember, Apply a Procedure, or Understand. This study has shown that questions asked in class by instructors teaching proof-based courses seldom ask Recall or Apply a Procedure questions (only 9.5% of the total number of questions when all instructor questions are compiled). These instructors asked more higher-order questions at this level (38.8% of the total number of instructor questions were at the level of Apply Understanding and above). This stands in contrast to similar results at the K-12 level. The difference in the types of questions may be because the mathematics content in proof-based courses is not procedural, but requires higher-order thinking.

This study did not take into account the effects of linked questions. It may be that the instructors begin with a higher-level question, but use a funneling or focusing pattern to reduce the cognitive level of the question (Fukawa-Connelly, 2012a). This question can be addressed by either doing an analysis of the questions that were actually answered by the students, or by doing an analysis of the question clusters that appeared in the data. Future research using this data will include a collaborative researcher who can code the data using the developed coding scheme to establish inter-rater reliability.

References

Anderson, L. & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Bagnato, R. (1973). Teaching college mathematics by question-and-answer. *Educational Studies in Mathematics, 5*, 185-192.

Bressoud, D. (2011). The calculus I instructor. *Launchings*, http://www.maa.org/columns/launchings/launchings_06_11.html

Epstein, J. (2007). Development and Validation of the Calculus Concept Inventory. In *Proceedings of the Ninth International Conference on Mathematics Education in a Global Community*, Palermo

Fraivillig, J., Murphy, L., & Fuson, K. (1999). Advancing children's mathematical thinking in Everyday Mathematics classrooms. *Journal for Research in Mathematics Education, 30*(2), 148-170.

Fukawa-Connelly, T. (2012a). A case study of one instructor's lecture based teaching of abstract algebra. *Educational Studies in Mathematics, 81,* 325-345.

Fukawa-Connelly, T. (2012b). Classroom sociomathematical norms for proof presentation in undergraduate in abstract algebra. *Journal of Mathematical Beahvior, 31*. 401-416.

Gall, M. (1970). The use of questions in teaching. *Review of Educational Research, 40*(5), 707-721.

Henningsen, M. & Stein, M. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524-549.

Krantz, S. (1999). *How to Teach Mathematics: Second Edition*. Providence, RI: American Mathematical Society.

Lobato, J., Clarke, D., & Ellis, A. (2005). Initiating and eliciting in teaching: A reformulation of telling. *Journal for Research in Mathematics Education, 36*(2) 101-136.

McClain, K. & Copp, P. (2001). An analysis of development of sociomathematical norms in one first grade classroom. *Journal for Research in Mathematics Education, 3*, 236-266.

McKeachie, W. & Svinicki, M. (2006). *McKeachie's Teaching Tips: Twelfth Edition.* New York: Houghton Mifflin Company.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Sahin, A. & Kulm, G. (2008). Sixth grade mathematics teachers' intentions and use of probing, guiding, and factual questions. *Journal of Mathematics Teacher Education, 11,* 221-241.

Sawada, D., Piburn, M., Judson, E., Turley, J., Falconer, K., Binford, 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.

Speer, N., Smith, J., & Horvath, A., (2010). College mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior, 29,* 99-114.

Stuessy, C. (2006). Mathematics and science classroom observation protocol system (M-SCOPS): Classroom observation and videotape analysis of classroom learning environments. A manual prepared for a 2-day workshop prepared for mentors of intern teachers in the PLC-MAP project: College Station, TX, Department of Teaching, Learning, and Culture.

Tallman, M. & Carlson, M. (2012). A characterization of calculus 1 final exams in U.S. colleges and universities. *The 15th annual conference on Research in Collegiate Mathematics Education*¸ Portland, OR.

Thompson, P., Castillo-Chavez, C., Culbertson, R., Flores, A., Greeley, R., Haag, S., Lawson, A., Rose, S., & Rutowski, R. (2007). Failing the future: Problems of persistence and retention in science, technology, engineering, and mathematics (STEM) majors at Arizona State Univsrsity. *Report Prepared & Submitted by the Provost's Freshman STEM Improvement Committee.*

Tobin, K. (1986). Effects of teacher wait time on discourse characteristics in mathematics and language arts classes. *American Educational Research Journal, 23* (2) 191-200.

van Zee, E. & Minstrell, J. (1997). Using questioning to guide student thinking. *The Journal of the Learning Sciences, 6*(2) 227-269.

Voigt, J. (1989). Social functions of routines and consequences for subject matter learning. *International Journal of Educational Research, 13*(6), 647-656.

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.

Wood, T. (1994). Patterns of interactions and the culture of mathematics classrooms. In Lerman, S. (ed). *Cultural Perspectives on the Mathematics Classroom.* Kluwer Academic Publishers.

Wood, T. (1999). Creating a context for argument in mathematics class. *Journal for Research in Mathematics Education, 30* (2), 171-191.

Appendix A

Bloom's Taxonomy Codes adapted to teaching advanced (proof-based) mathematics.

Remembering	Applying a Procedure	Understanding	Applying Understanding	Analyzing	Evaluating	Creating
Retrieve relevant knowledge from long-term memory Recall a familiar definition, formula, result, or proof technique Cite Identify Recall List	Students must recognize what knowledge or procedures to recall when directly prompted to do so. Perform computations in a familiar context Using a procedure that does not require deep understanding Plug-in numbers to a familiar function Compute the base case for an induction proof	Make interpretations, provide explanations, make comparisons, or make inferences that require understanding of a mathematics concept Explain in your own words Interpreting definitions Explain the implications of a condition or definition Understand and explain the steps in a given proof Explain, interpret, or modify notation Describe the elements of a given set Describe the graph of a given function Classifying	Recognize when to use or apply a concept Applying a procedure in an unfamiliar context Deciding which theorem or definition to apply and using it. Applying a theorem, result, or proof technique to an unfamiliar context. Plug-in numbers to an unfamiliar function (prime counting function) Computing in an unfamiliar context (different bases) Identifying properties of a certain function	Break material into parts and determine how they relate. Examine, organize, generalize, differentiate Detecting a contradiction Determining the next step in a proof (when it is more than just interpreting a definition) Deconstruct a statement into a plan for a proof Focusing Clarifying the statement to be proved Breaking down the proof into parts Determining how the elements fit or function within a structure	Make judgments based on criteria and standards. Conclude, testing, justify, proving, validate, defend, assess Give a more efficient computation Which proof method is best? Justify your thinking Justify a step in a proof Explain why a certain method was used.	Put elements together to form a coherent or functional whole Reorganize elements into a new pattern or structure. Hypothesizing Designing Constructing Construct a set or function that has certain properties Conjecture Formulate a new definition

COVARIATIONAL REASONING AND GRAPHING IN POLAR COORDINATES

Kevin C. Moore, Teo Paoletti, Jackie Gammaro, and Stacy Musgrave
University of Georgia

Researchers continue to emphasize the importance of covariational reasoning in the context of students' function concept, particularly when graphing in the Cartesian coordinate system (CCS). In this manuscript, we extend this body of literature by characterizing two pre-service teachers' thinking during a teaching experiment focused on graphing in the polar coordinate system (PCS). We illustrate how the participants engaged in covariational reasoning to make sense of graphing in the PCS and make connections with graphing in the CCS. By foregrounding covariational relationships, the students came to understand graphs in different coordinate systems as representative of the same relationship despite differences in the perceptual features of these graphs. In synthesizing the students' activity, we provide remarks on instructional approaches to graphing and how the PCS forms a potential context for promoting covariational reasoning.

Key words: Polar coordinates, Covariational reasoning, Graphing, Function, Multiple representations

First introduced at the elementary level, graphs are essential representations for the study of numerous mathematical topics including modeling relationships between quantities, exploring characteristics of functions, and the study of analysis. Reflecting the heavy focus on graphing in school mathematics, mathematics education researchers have given significant attention to graphing, with a multitude of researchers (see Carlson [1998] and Oehrtman, Carlson, & Thompson [2008] for review) having investigated students' graphing meanings in the context of function. Although graphing receives significant attention in mathematics education research, little of this focus has been given to graphing in the *polar coordinate system* (PCS). Complicating the matter, available research on students' meanings for the PCS (Montiel, Vidakovic, & Kabael, 2008; Montiel, Wilhelmi, Vidakovic, & Elstak, 2009; Sayre & Wittman, 2007) emphasize student difficulties, with these difficulties often involving problematic connections with the *Cartesian coordinate system* (CCS).

In the present study, we characterize student thinking when graphing in the PCS and draw connections with existing research on graphing and function. Specifically, we discuss two undergraduate students' reasoning when graphing in the PCS. To graph relationships in the PCS, both students engaged in several ways of thinking that ranged from determining and plotting discrete points to reasoning about how quantities continuously vary in tandem. The former way of thinking enabled them to gain a sense of more basic functions (e.g., constant rate of change relationships), but was not sufficient when graphing more complex functions (e.g., trigonometric relationships). The latter way of thinking enabled the students to flexibly graph relationships in the PCS by systematically generating curves as representative of a dynamic relationship between two quantities' values. Collectively, the students' ways of thinking enabled them to approach graphing in the PCS and CCS in the same manner despite differences in the coordinate systems' conventions. Against the backdrop of these findings, we discuss how the PCS offers a *potential* setting for promoting covariational reasoning.

Background

Often first introduced in a precalculus course, the PCS is critical for the study of advanced mathematics and can be found throughout STEM fields. For instance, in complex analysis the operations of multiplication, division, and exponentiation are more readily explored when using the polar form of complex numbers. Although the PCS is critical for the aforementioned areas, research on student thinking in the context of the PCS is sparse, with a pair of studies by Montiel and colleagues (Montiel et al., 2008; Montiel et al., 2009) forming the most applicable works to the present study. Both studies included a focus on the PCS, with the earlier study (Montiel et al., 2008) exploring relationships among two-dimensional coordinate systems and the subsequent study (Montiel et al., 2009) including two- and three-dimensional coordinate systems.

Of relevance to the present study, Montiel et al. (2008) identified that the connections students create between the CCS and PCS are tied to CCS conventions. For instance, students in their study often relied on rules learned in relation to the CCS to determine if a given relation is a function, which included applying the vertical line test to determine if a graphed relationship in the PCS is a function. Some students also referenced "known" functions when determining if graphs were functions. By "known", we interpret the authors to mean that the students recalled a shape in the plane that they had previously deemed a function (e.g., a student claiming that an upside down parabola is a function because parabola's are defined as such). Like the 2008 Montiel et al. study, Montiel et al. (2009) noted that when students moved among representational systems, the students' function meanings did not include coordinating different conventions of the representational systems, with the students often relying on the conventions from one representational system (e.g., the CCS).

Montiel and colleagues' findings (Montiel et al., 2008; Montiel et al., 2009) relative to students' difficulty in coordinating representational systems speak to Lobato and Bowers's (2000) and Thompson's (1994c; in press) statements about multiple representations. Lobato and Bowers (2000) questioned, "…whether tables, graphs, and equations are multiple representations of *anything* to students" (p. 4). When framing multiple representations in terms of connected areas of representational activity, Thompson (1994c) explained:

> I agree with Kaput (1993) that it may be wrongheaded to focus on graphs, expressions, or tables as representations of function, but instead focus on them as representations of something that, from the students' perspective, is *representable,* such as some aspect of a specific situation. The key issue then becomes twofold: (1) To find situations that are sufficiently propitious for engendering multitudes of representational activity and (2) Orient students to draw connections among their representational activities in regard to the situation that engendered them. (p. 39-40)

Moreover, if students are to conceive multiple representations of *something*, then it is necessary that they not only construct the *something* that is to be represented, but also have meanings for the representational systems such that when the students operate within and move among systems, they can think about their representational activity as conveying the same *something*. Returning to the studies by Montiel and colleagues, the students did not appear to have distinct meanings for the coordinate systems that simultaneously supported connections among these systems.

Theories of quantitative and covariational reasoning (e.g., Carlson, Jacobs, Coe, Larsen, & Hsu, 2002; Thompson, 1990, 2011) foreground students' construction of what Thompson identified as "the something that, from the students' perspective, is *representable*." Quantitative reasoning (Thompson, 1990) emphasizes students' construction of quantities and relationships among quantities in such a way that the resulting quantitative structure forms the foundation for their mathematical activity. An aspect of quantitative reasoning is conceiving how two quantities vary in tandem, termed covariational reasoning (Carlson et al., 2002). Quantitative and covariational reasoning are central to students' understanding of numerous undergraduate mathematics topics, including exponential relationships (Castillo-Garsow, 2012; Confrey & Smith, 1995), trigonometric relationships (Moore, 2012), rate of change (Carlson et al., 2002; Thompson, 1994a), function (Oehrtman et al., 2008), and the fundamental theorem of calculus (Thompson, 1994b).

In characterizing second-semester calculus students' thinking, Carlson et al. (2002) identified how the students' covariational reasoning influenced their ability to make sense of dynamic situations and graphing. Specifically, the authors identified several mental actions that the students engaged in when coordinating how quantities vary in tandem. These mental actions included, but were not limited to, coordinating directional change (e.g., quantity A increases while quantity B increases), coordinating amounts of change (e.g., the increase in quantity A increases for successive increases in quantity B), and coordinating rates of change (quantity A increases at an increasing rate with respect to an increasing quantity B).

The tasks found in Carlson and colleagues' (2002) study involved the CCS, but the mental actions associated with covariational reasoning need not be specific to the coordinate system in which one is graphing (nor are they specific to the act of graphing). To illustrate, consider graphing the relationship defined by $f(b)=b^2$. For the function f and $b > 0$, as the *input* increases, the *output* increases at an increasing rate. It follows that the *output* increases such that the *change of output* increases for successive equal *changes of input*. These *change of output* values increase by a constant amount for successive equal *changes of input*; as the *input changes* from 0 to 1 to 2 to 3 and so on, the *output increases by* 1, 3, 5, and so on. The aforementioned covariational relationship can be represented in both the CCS ($y=x^2$, Figure 1, left) and PCS ($r = \theta^2$, Figure 1, right). Changing coordinate systems results in a different pictorial object, but covariational reasoning enables conceiving the graphs as one in the same; changing the coordinate system changes the coordinate conventions and the shape of the curve, but the defined covariational relationship remains invariant. It is in this way that the graphs represent the same *something*–a structure of related quantities.

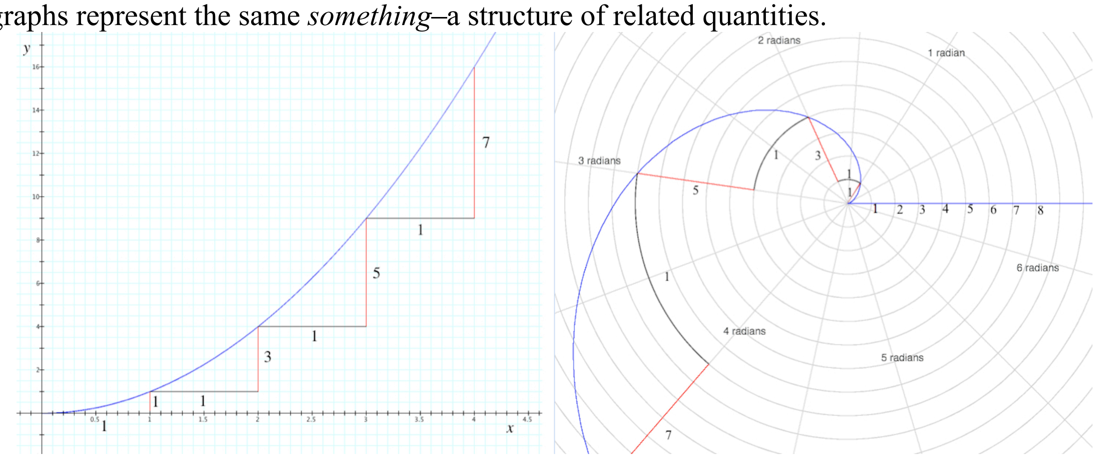


Figure 1. Representing a covariational relationship: $y=x^2$ (left) and $r = \theta^2$ (right).

We note that another common topic is determining both the PCS and CCS mappings that define a specified curve in a plane (e.g., Figure 2). In such a situation, the invariant object is a curve in the plane, and the curve defines different relationships depending on the coordinate system of choice. This differs from the prior discussion that focused on a covariational relationship as the invariant object (e.g., Figure 1). When tasked with determining if $r = 2$ defines a function, one student in the Montiel et al. (2008) study first determined the CCS representation of the curve defined by $r=2$, namely $x^2+y^2=4$. The student then used the CCS representation and formula to claim that $r=2$ is not a function. While the authors presented both a curve and formula in their problem, if only presented with a curve in the plane (e.g., Figure 2) and asked to determine if the curve defines a function, we might expect that particular student to respond in a similar way (e.g., basing his answer on the CCS to claim that the curve does not define a function). Reconciling this problem requires that the student understand the ambiguous nature of questioning whether a curve in a plane defines a function: the answer depends upon the coordinate system of choice, as different coordinate systems correspond to different relationships. It is not the curve that is or is not a function, but instead the relationship defined by the curve, and this relationship is dependent on the chosen coordinate system (Figure 2).

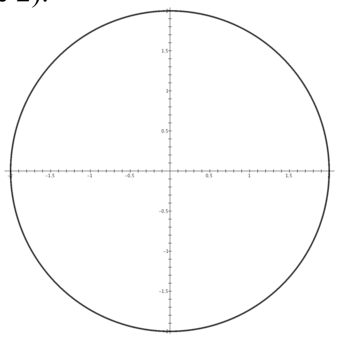


Figure 2. A curve defined by $r = 2$, $x^2 + y^2 = 4$, and $(x(t), y(t)) = (2\cos(t), 2\sin(t))$.

Methodology

In the present study, we conducted a teaching experiment with the intention of characterizing models of students' mathematics (Steffe & Thompson, 2000). The following research questions guided the study: (i) what ways of thinking do students engage in when graphing in the PCS? and (ii) how do students use the CCS when graphing in the PCS?

Subjects and Setting

The subjects of this study, 'John' and 'Katie', were two undergraduate students enrolled in a pre-service secondary mathematics education program at a large public university in the southeast United States. At the time of data collection, the participants were third year students (by credits taken) and taking the first pair of courses (one methods and one content) in their program. We chose the students on a voluntary basis from the content course.

We chose to work with pre-service teachers for several reasons. First, the present study was situated within a series of studies (e.g., Moore, 2012, in press) into students' and teachers' quantitative reasoning in the context of trigonometric functions and the PCS. The content course from which the students volunteered covered trigonometric functions and graphing in multiple coordinate systems, and thus the students were a natural fit for the study. Second, our previous experiences teaching the course suggested that students' ways of thinking about graphing became problematic when attempting to graph relationships in multiple coordinate systems. Thus, we intended to gain insights into these difficulties and the ways of thinking that help or inhibit students' ability to reconcile such difficulties.

The content course engaged the students in quantitative reasoning and covariational reasoning to explore topics central to secondary mathematics (e.g., trigonometry, exponential functions, linear functions, rate of change, and accumulation). Prior to the present study, the course explored ideas of angle measure and trigonometric functions. The approach to these topics was grounded in previous research (Moore, 2012, in press) on students' learning of angle measure and trigonometric functions. Due to the focus on covariational reasoning in the context of trigonometric functions, we expected the students to be familiar with covariational reasoning when entering the study. Based on previous research highlighting student difficulties with the PCS (Montiel et al., 2008; Montiel et al., 2009) and covariational reasoning (Carlson et al., 2002), we questioned whether they would or would not spontaneously engage in said reasoning when graphing in the PCS.

Data Collection and Analysis

The teaching experiment (Steffe & Thompson, 2000) consisted of five 75-minute teaching sessions with the student pair. The first teaching session developed conventions of the PCS (e.g., coordinate pairs representing the distance from a fixed point and the radian measure of an arc) and supported students' spatial reasoning in the PCS (e.g., considering the location of a point that has a varying arc measure and a constant distance measure, and vice versa). The subsequent teaching experiment sessions, which are the focus of the present report, involved graphing functions of the form $r = f(\theta)$ or $\theta = g(r)$ in the PCS.

Data collection involved videotaping all teaching sessions and student work. Also, fellow researchers observed each teaching session and took notes on the interactions between the researcher and students. We debriefed immediately after each session in order to discuss the students' thinking and document all instructional decisions. Our retrospective analysis of the data involved first transcribing the entire data set and then identifying instances offering insights into the students' thinking. We then performed a conceptual analysis (Thompson, 2000) of these instances in order to generate and test models of the students' thinking so that these models provided viable explanations of their behaviors. We particularly sought to characterize the students' reasoning when graphing in the PCS and CCS.

Results

As described in the prior section, we first worked with the students to support their construction of the PCS.[1] Nearly two full sessions were dedicated to exploring the conventions of the PCS including: (i) the quantitative basis for the coordinate system (e.g., a directed radial distance and a directed angle measure); (ii) each point in the plane can be represented by an infinite number of coordinate pairs (e.g., $(r,\ \theta \pm 2n\pi)$ defines the same point in the plane for all integers n); and (iii) how variations in one coordinate value influence the location of a point in the plane. Of relevance to the present study, the students' actions did not suggest familiarity with the PCS. The students verbalized that they had "seen" the PCS in previous courses, but that they had not explored the aforementioned features of the PCS. Following the introduction to the PCS, we transitioned into graphing relationships including linear, quadratic, and trigonometric relationships.

Graphing a Linear Relationship

We first tasked the students with graphing $f(\theta) = 2\theta + 1$, with an interest in how the students would reconcile a linear relationship not being represented by a line in the PCS. John and Katie initially graphed the relationship in the CCS (e.g., $y = 2x + 1$) by identifying both the x- and y-intercepts and connecting these points with a line. This method did not carry over to the PCS, and to gain a better sense of the function they plotted points for θ values of 0, 1, 2, 3, and 4. They then connected these points (Figure 3) and related the two graphs (Table 1).

Table 1

John and Katie discuss two "linear" graphs

Katie:	They went out by two, like you know here (*pointing at the two in the formula* $r = 2\theta + 1$) the slope is like two (*tapping along the CCS graph*).
Int.:	This has no slope (*pointing to the PCS graph*)…
Katie:	No, I'm relating the slope here (*pointing to the CCS graph*), to the difference in the radius of two each time (*tapping along the PCS graph*). Like [the radius is] one, three, five, seven, nine, eleven (*pointing to the corresponding points on the polar graph*), [the radius] increases by two.

Katie reasoned that both graphs convey some quantity changing by amounts of two for successive changes of one in the other quantity. In the case of the CCS, she related this

[1] These sessions are the focus of a manuscript under preparation.

feature to "slope," meaning the amount by which y increased for successive equal changes in x. In the case of the PCS, she illustrated the increase in the distance from the pole, which she referred to as the "radius," for successive changes of angle measure.

As the interaction continued, John and Katie denoted the aforementioned amounts of change (Figure 3) and claimed that both graphs convey a "constant rate of change" between the input and output values. Katie then added, "That's cool…because you'd never see this (*referring to the PCS graph*) and be like, that's a linear function." Katie's comment underscores an important aspect of the pair's thinking. From a shape standpoint, the PCS graph is not "linear". But, by considering the graphs in terms of covarying quantities, they conceived both graphs as representative of the same relationship (e.g., a relationship such that there is a constant rate of change between the quantities).

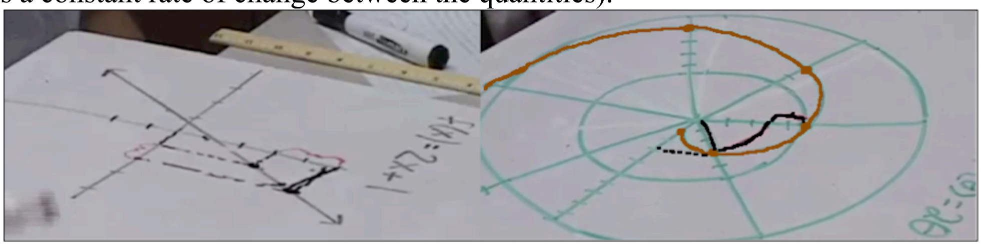


Figure 3. Students' graphs of a "linear" relationship.

Graphing a Quadratic Relationship

Based on John and Katie's approach to graphing a linear relationship, we conjectured that similar reasoning would enable them to graph a relationship with a non-constant rate of change. We asked the students to graph $r = \theta^2$, which they compared to $r = \theta$. Like their activity when graphing the linear relationship, the students first plotted points and connected the points in both the CCS and PCS. They then compared their graphs (Figure 4, Table 2).

Table 2

John and Katie discuss a quadratic relationship

Katie:	r of theta, but compared to r of theta squared, [r of theta squared is] like expanded (*Katie points to the two graphs and then spreads her hands apart*). Like, like, this one's (*referring to linear graph*) like much more tighter swirled (*moving her hands in a circular motion*) but then this one (*referring to quadratic graph*) is just like looser I guess.
John:	Yeah, we can see better, with both of them, both graphs, that the change in radius (*referring to quadratic*) for every radian further that the angle is increasing (*rotating his hand in successive rotations while spreading his index and middle finger apart*)…Um, the radius, every time is increasing at an increasing rate (*referring to quadratic*).
Int.:	Okay now what's that mean in terms of amounts of change?
John:	We could do equal changes in theta and then…
Katie:	Like, if we looked at first these two then these two points (*indicating the points (9, 3) to (16, 4), and then (16, 4) to (25, 5)*), the change of theta here would be this, that length (*drawing an arc from (9, 3) to (9, 4)*). But then the change is radius would be up that line (*drawing a segment from the point (9, 4) to (16, 4)*).
John:	Which is seven.
Katie:	And then we have the same thing (*draws an arc from (16, 4) to (16, 5) and a segment from (16, 5) to (25, 5)*)…so you can see these black lines, the [change in radius] is increasing.
John:	So that's like nine to sixteen (*pointing to the segment connecting the points*

23		*(9, 4) and (16, 4)*), which is seven, and this one is sixteen to twenty-five
24		(*pointing to the segment connecting the points (16, 5) and (25, 5)*), which is
25		nine, which we can see there too (*pointing to the CCS graph*).

Katie first compared the perceptual shapes of the graphs, observing that the quadratic is "expanded" (lines 1-5). Such thinking is similar to a student observing one graph to be "steeper" or growing "faster" than another graph or considering one graph to be the result of "stretching" another graph in the CCS. As the interaction continued, the students' actions suggested that their thinking was not constrained to thinking about the curves solely as pictorial objects with comparable qualitative features. The students worked together to describe the behavior of the graphs in terms of amounts of change and rates of change between the two quantities. Specifically, the students reasoned that the graph of $r = \theta^2$ is "looser" or moves away from the pole "faster" because r increases at an increasing rate with respect to θ, which they confirmed by identifying specific changes in the quantities' values (lines 13-25).

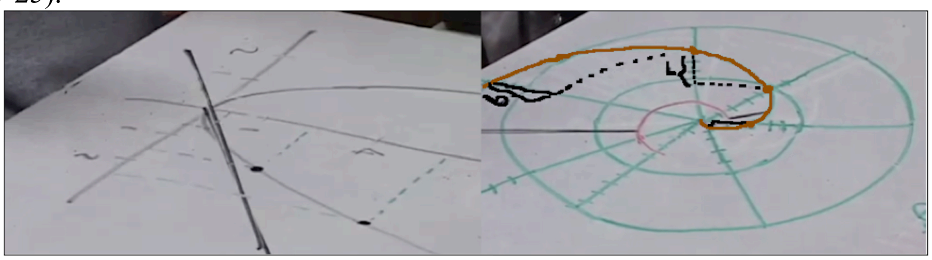


Figure 4. Students' graphs of the quadratic relationship.

Immediately following John's last statement (line 25), the students denoted amounts of change on a graph in the CCS (Figure 4), while Katie claimed, "Like our change in input here (*referring to CCS graph*) would represent the change in this angle measure (*indicating the corresponding change on the PCS graph*), and then our output, change of radius length, and that's increasing for equal changes." Compatible with their actions when graphing a linear relationship, John and Katie conceived both graphs as a structure of covarying quantities, which enabled them to understand a graph in the PCS and a graph in the CCS as one in the same; to John and Katie, both graphs represented the same covariational relationship.

Conceiving a Trigonometric Relationship

During the first two tasks, John and Katie revealed a tendency to engage in covariational reasoning when graphing in both coordinate systems. We questioned whether they would engage in such reasoning to interpret given graphs. As Montiel et al. (2008) showed, students have a tendency to rely on the CCS representations of given curves in a plane.

We tasked the students with determining a formula for a graph of the sine function ($r = \sin(\theta)$, Figure 5). We chose the sine function to provide a curve that differed in shape from those they had explored during previous tasks, particularly in that the relationship is not monotonic for $\theta > 0$. Like the previous problems, John and Katie first focused on identifying specified coordinate pairs. After identifying r-values corresponding to θ values of 0, π/2, π, and 3π/2, the students conjectured that $r = \sin(\theta)$ is the appropriate formula for the given graph. The students then drew a graph of the sine function in the CCS and justified their conjecture (Table 3).

Table 3

John and Katie discuss a trigonometric relationship

1	Katie:	So we start here (*pointing to the pole in the PCS*).
2	John:	Ya, and we're sweeping around (*making a circular motion with his hands*).
3		As theta's increasing, distance away from the origin is increasing (*Katie*

traces along the polar graph from 0 radians to π/2 radians) and then decreases again…it increases until pi-over-two and then it starts decreasing.

Int.: And then what happens from like pi to two-pi?

Katie: It's the same.

John: Same idea except, the radius is going to be negative, so it gets more in the negative direction of the angle we're sweeping (*using marker to sweep out a ray from π to 3π/2 radians – see Figure 5*) until three-pi-over-two, where it's negative one away. And then it gets closer.

Int: OK, so from like three pi over two to two pi, could you show me where on this graph we'd start from and end at.

Katie: This is the biggest in magnitude, so it's the furthest away (*placing a finger on a ray defining 3π/2 and a finger at (1, π/2)*), and then [the distance from the pole] gets smaller in magnitude (*simultaneously tracing one index finger along an arc from 3π/2 to 2π and the other index finger along the graph – see Figure 5*).

When making sense of the graph, and compatible with the previous interaction (Tables 1-2), the students used a combination of identifying discrete points and reasoning about the quantities covarying between these points. When working between points, the students reasoned that the graph conveyed the distance from the pole as increasing or decreasing for a continuously increasing angle measure (e.g., lines 2-5 & 14-18). For instance, Katie reasoned that as the angle measure increases from 3π/2 radians to 2π radians, the distance from the pole decreases from a magnitude of one to zero (a value of -1 to 0) (Figure 5).

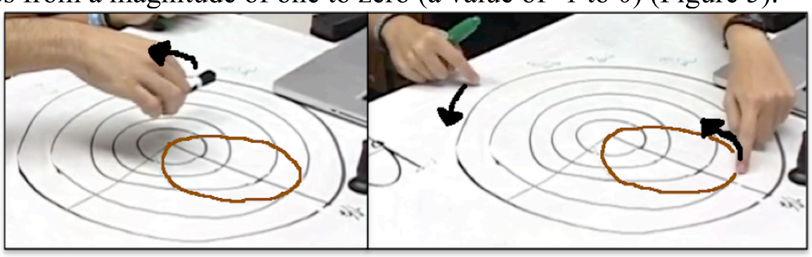


Figure 5. Illustrating covarying quantities.

Stemming from instruction occurring earlier in the semester, the students understood the sine function to convey such a covariational relationship and thus concluded that the graph represents the relationship defined by $r = \sin(\theta)$. Following this interaction the students justified their formula by identifying that the graph of $y = \sin(x)$ in the CCS conveys the same covariational relationship and has identical critical points as the PCS graph. To the students, the CCS and PCS graphs of the sine function represented the same covariational relationship.

Graphing a Trigonometric Relationship

To conclude the teaching experiment, we tasked John and Katie with graphing the relationship defined by $r = \sin(2\theta)$. Contrary to their activity on previous tasks, they first attempted to think about the graph in terms of a transformation of $r = \sin(\theta)$. Katie made motions to indicate the CCS graph of $y = \sin(x)$ stretching to an amplitude of two and the PCS graph changing such that the maximum distance from the pole is two. John did not agree with Katie, arguing that the radial value would be 0 for $\theta = \pi/2$ as opposed to two, and thus the "function reaches zero again at pi over two, so instead of having the circle go until pi radians, the circle would complete at pi over two." Like Katie, John imagined transforming the circle defined by $r = \sin(\theta)$, but in his case he imagined transforming the PCS graph such that the circle was contained between the rays defining the angle measure of π/2 radians.

Despite their focus on transforming a known graph, the students did not sketch a new graph. Instead, they determined coordinate pairs, choosing to use successive increments of $\pi/4$ for θ. As they determined the r-values, John expressed, "we may not actually get shapes," and then refined their increments to intervals of $\pi/8$ for θ. His actions suggest that he was concerned with how the two quantities covaried within the successive $\pi/4$ intervals and believed they should consider variation within these intervals before completing their graph.

After the students determined the *r-values* for successive values of θ that differed by $\pi/8$, they plotted the points and connected these points (Figure 6). The students were then asked to describe the first "petal" of the graph (Table 4).

Table 4

John and Katie justify a petal on their graph

Int.:	What's it conveying about the relationship between r and theta for that petal?
Katie:	So the relationship is going to reverse at pi over four. Right? Cause it's like…
John:	Well, yeah cause if we look at our, at pi over two (*tracing ray*), er, pi over two, on the graph of, oh, was it r equals sine theta.
Katie:	Well like here, r is increasing (*tracing along graph from 0 to π/4 radians*) as theta increases (*tracing along circle from 0 to π/4 radians*). But then as theta is increasing (*tracing along circle defining from π/4 to π/2 radians*) r is decreasing (*tracing along graph from π/4 to π/2 radians*).
John:	Yeah, like it does at pi over two for sine of theta (*tracing* $r = \sin(\theta)$ *in the air*).

Katie conceived the first petal as conveying that r increased and then decreased as θ continuously increased from 0 to $\pi/2$ radians. In trying to relate this relationship to a graph of $r = \sin(\theta)$, John continued on to describe that both graphs have the same pattern, but over different input intervals. Katie then added, "Oh, you're saying on this graph, like that (*drawing sine curve in Cartesian plane*), then r increases (*tracing increasing part of the curve*) then r decreases (*tracing decreasing part of the curve*)." John continued by drawing a graph of $r = \sin(2\theta)$ in the CCS to show that both $r = \sin(2\theta)$ and $r = \sin(\theta)$ convey an increasing, decreasing, increasing, decreasing, and so on pattern, to which the interviewer asked them to explain how the graphs of the posed relationship ($r = \sin(2\theta)$) related (Table 5).

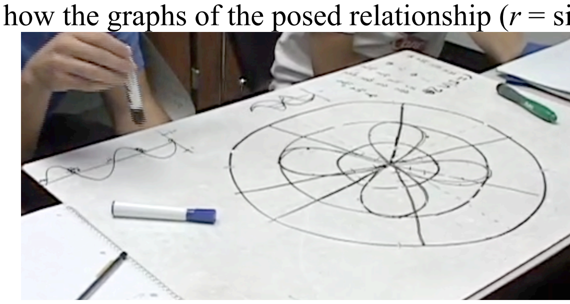


Figure 6. John and Katie's board work when graphing $r = \sin(2\theta)$.

Table 5

John and Katie relate two graphs from different systems

John:	Um, well at, at pi over four, the value of this graph is one (*identifying point on the CCS graph*), which is kinda how we see, when…
Katie:	The radius (*horizontally moving her finger from pole to a radius of 1*) of one (*moving her finger along the arc to the point (1, π/4)*).
John:	This is pi over four…
Int.:	Okay.
John:	…the angle measure, and that's one, the radius (*identifying point on the PCS graph*). And then at pi over two the angle measure, we have zero radius here

	(*identifying points on both graphs*), it returns to zero (*tracing PCS graph from previous point*).
Int.:	Okay.
Katie:	And then you can tell how this increased, like, you can see, by the shape of the graph, makes sense, of the like the curve (*tracing her hand along the two curves throughout description*).
Int.:	Okay so say a little bit more about that. Why does that make sense? What is that graph showing that this graph is also showing?
Katie:	Well like here, *r* is increasing (*tracing CCS graph*). Like to pi over [four] (*pointing to π/4 on horizontal axis*), we see *r* is increasing here (*motioning to PCS graph*), but it's also increasing at a decreasing rate (*motioning to CCS graph*). So like here (*motioning to PCS graph*), *r* gets, the changes in *r* between these points is, like here (*identifying points at successive angle measures between 0 and π/4*), like the change in *r* is really decreasing. But then this is like the reverse here, it's decreasing at an increasing rate (*motioning to both graphs*).

Katie continued discussing a covariational relationship as she considered the two graphs, suggesting that she understood both graphs as conveying the same relationship. We note that this interaction also demonstrates a shift in the students' roles when compared to their activity during previous interactions. Whereas Katie typically first plotted points and then John reasoned about the covariational relationship between these points, this interaction reveals John making the move to identify points when comparing the two graphs (lines 1-2 & 7-10). Katie then described the covariation between these points in a smooth manner (lines 12-24). Also, throughout the interaction and compatible with the students' actions during previous interactions, Katie made smooth motions with her hands as she was describing information about the two graphs. Likewise, as the discussion continued after the interaction in Table 5, the students simultaneously traced along corresponding parts of both graphs while describing the covariational relationship (e.g., "*r* decreases at an increasing rate"), further illustrating the students' consideration of both graphs as representative of the same relationship.

Discussion and Implications

During each task, John and Katie's activity foregrounded covariational relationships and the students conceived graphs in both the CCS and PCS as representative of these relationships. Importantly, the students' ways of thinking were not constrained to one coordinate system, nor did the students appear to rely on one coordinate system over the other. In this section, we discuss various aspects of the students' ways of thinking and then describe implications of these ways of thinking in the context of the PCS, covariational reasoning, and school mathematics.

Coordinating Images of Change

The various ways in which students make sense of covarying quantities are complex and have serious implications for student learning (e.g., Carlson et al., 2002; Confrey & Smith, 1995; Thompson, 1994a, 1994b). Of relevance to the present study, researchers (Castillo-Garsow, 2012; Castillo-Garsow, Johnson, & Moore, submitted) have characterized two different images of change: *chunky* and *smooth*. As Castillo-Garsow and colleagues (2012, submitted) described, chunky images of change entail a nearly singular focus on discrete points, which the authors likened to considering how the volume and height of water in a bottle varies when filling the bottle cup by cup. Smooth images of change, however, entail a continuous change in progress, which the authors likened to considering how the volume and height of water varies when filling up a bottle of water by turning on a water spigot.

In John and Katie's case, they typically relied on first graphing discrete points and comparing discrete amounts of change between these points, which is a prototypical action of chunky thinkers. Differing from these chunky-like actions, the students exhibited behaviors that are suggestive of considering continuous change within these chunks, thus illustrating that they were not constrained to thinking about discrete points and comparing discrete amounts of change. For instance, to make sense of curves between identified points the students imagined one quantity increasing or decreasing for a continuous increase in the other quantity (e.g., Tables 3-5). Such thinking was particularly apparent when the students graphed $r = \sin(2\theta)$. During their progress on this task, John found it important (e.g., "we may not actually get shapes") to consider variation between already identified coordinate pairs. John's concern prompted the students to refine their intervals and subsequently focus on smooth variation.

Collectively, the students' actions highlight the importance of both chunky and smooth images of change for graphing relationships and constructing connections among coordinate systems. Even though the students sometimes held a focus on discrete points and amounts of change, their subsequent decisions and explanations were often informed by their awareness that quantities covaried continuously between points. By coordinating these images of change in different coordinate systems, the students foregrounded covariational relationships when operating in various representational systems. Most importantly, making sense of these covariational relationships through various images of change enabled them to move among representations–both in coordinate systems and formulas–while holding in mind something invariant about these representations. It was in this way that the students saw different representations as one in the same. Each representation represented something *representable*: a covariational relationship that entailed a structure of related quantities.

John and Katie's coordination of different images of change is consistent with that of *continuous covariation* (Thompson, 2011, in press), which can be framed as a combination of smooth and chunky thinking (Castillo-Garsow, 2012; Castillo-Garsow, Johnson, & Moore, submitted). An essential aspect of continuous covariation is that any conceived chunk of variation (e.g., the interval between two discrete points) is understood as entailing smooth variation, where this smooth interval of variation can be partitioned into sub-chunks. These sub-chunks are also understood as entailing smooth variation that can be partitioned, and so on (Castillo-Garsow, 2012; Thompson, 2011, in press). An important feature of continuous covariation, which is illustrated by John and Katie's actions, is that it involves more than a pointwise understanding of graphing or function. For instance, John and Katie not only understood that graphs in different coordinate systems involved equivalent coordinate pairs, but they also conceived invariance among the graphs in terms of a dynamic and continuous relationship between quantities' values. The students conceived graphs in terms of both discrete points and two dynamically changing quantities' values, and considering smooth variation between discrete points was central to their making sense of and producing graphs.

Polar Coordinates and School Mathematics

Returning to Thompson's (1994c) discussion of multiple representations, he emphasized supporting students in conceiving something representable, finding situations that students are likely to use multiple representations, and directing students to make connections among their representational activity. John and Katie's activity indicates that engaging in covariational reasoning supported them in conceiving something that is representable, namely a dynamic relationship between two quantities. Also in line with Thompson's claim, graphing in the PCS seems to have offered John and Katie situations that engendered multiple representations and directed them to make connections among their representational activity.

The body of literature on covariational reasoning suggests that US students typically do not receive repeated opportunities for covariational reasoning, nor do they experience a mathematics in which covariational reasoning is a fundamental activity (Carlson et al., 2002; Thompson, in press). As a result, students develop meanings for graphing and function that do not entail images of variation. For instance, Thompson (1994c), Monk (1992), and Goldenberg, Lewis, and O'Keefe (1992) documented that students are often oriented to interpreting graphs in terms of pictorial objects, as opposed to interpreting graphs as relationships between two quantities' values. Thinking of the nature identified by these researchers would lead a student to conclude that the two graphs in Figure 1 represent different relationships, as the graphs are perceptually different.

Students' propensity to treat graphs as pictorial objects is understandable if these students had experiences that predominantly involved one coordinate system; with one coordinate system receiving focus, viewing graphs as pictorial objects solves many of the school problems that students encounter. To illustrate, consider an approach to quadratic functions that first presents the rule of the function (e.g., $f(x)=ax^2+bx+c$) and then presents a graph (or graphs) to define the prototypical shape(s) of that function class (e.g., a parabola). From here, the approach might move into plotting points, defining properties of that shape (e.g., the vertex of a parabola), and exploring these properties in terms of transformations of other members of that function class (e.g., considering the vertex in terms of translating another vertex, or how $g(x)=f(x)+2$ can be thought of as a translation of $f(x)$). Such an approach, which is commonplace in US secondary and undergraduate precalculus, emphasizes graphs as pictorial objects made for manipulation. In the event that students do not already understand graphs as representative of covariational relationships, it can be expected that the student will construct function meanings inherently involving the manipulation of shapes in the plane, where these manipulations must adhere to some set of rules (e.g., $g(x)=f(x)+2$ means move the graph "up by 2"); if a student's activity is dominated by transforming shapes, then his abstracted meanings will likely be tied to such activity.

Students that develop graph and function meanings rooted in shapes and operations on these shapes are posited to solve problems set within the system in which those meanings were developed. But, when moving to a different system, meanings that entail activity dependent on system conventions become problematic because shapes and the representational conventions are changed. For instance, if a student relies on the vertical line test as in the example provided by Montiel et al. (2008), then a circle in the plane is not a function even when defined by a rule in polar form. Or, for a student who interprets the formula $y = x + 2$ as *line* or *constant slope* (slope meaning tilt), graphing $r = \theta + 2$ in the PCS can become problematic because the graph is not a line with constant slope (in the tilt sense).

In Katie's case, conceiving the PCS graph as a linear relationship was not problematic because her reasoning included considering how two quantities covary. This enabled her to conclude, "That's cool…because you'd never see this (*referring to the PCS graph*) and be like, that's a linear function." Likewise, both students' activity throughout the presented interactions (Tables 1-5) illustrates how covariational reasoning can give common meaning to different representational systems; covariational reasoning enabled seeing *something* as the same among their representational activity in multiple coordinate systems and was thus beneficial to the tasks at hand. Such activity illustrates that a *potential* benefit of incorporating the PCS in the study of mathematics is generating a context in which covariational reasoning can be engendered and emphasized as a beneficial way of thinking. By prompting students to consider multiple coordinate systems, a need can be established for ways of thinking (e.g., covariational reasoning) that enable approaching graphing in each system in a compatible way. By focusing on these ways of thinking, graphing in both systems might foster abstractions stemming from various operations involved in covariational reasoning. For instance, for Katie, graphing in both coordinate systems seemed to foreground the "constant rate of change" of a linear relationship (Table 1), as opposed to the visual tilt of a line.

A this time, the potential use of the PCS to promote covariational reasoning is little more than a conjecture on our part, but our conjecture does have foundations in learning theory, and particularly that of abstraction (Piaget, 1977). As Oehrtman (2008) characterized when discussing abstraction and instruction, abstraction theory has the following implications: (i) students' activity should reflect the structure of intended understandings; (ii) actions should be repeated and coordinated in a reflective way; and (iii) the development of a concept continually evolves through the attempted application of that concept or way of thinking in new situations. It is possible that infusing the PCS into instructional approaches to graphing can support a focus on covariational reasoning that is consistent with such a theory.

Returning to the situation in which students' graphing experiences lead to meanings that are tied to activity dependent upon the representational conventions (e.g., treating curves as pictorial objects for manipulation), it is likely that they will face perturbations when moving to the PCS (e.g., a linear relationship is not represented by a line in the PCS), thus creating a need for alternative ways of thinking about graphs in both the CCS and PCS. Covariational reasoning offers one such a way of thinking, and in such a case has the potential to change one's understanding of graphing in CCS while also supporting graphing in the PCS. In the case that a student already predominantly engages in covariational reasoning when operating in the CCS, then introducing the PCS will provide a situation in which they can repeat and coordinate these actions. In both cases, including multiple coordinate systems provides an opportunity to engage students in mutliple representational activities and support their abstraction of meanings rooted in the connections that they make among these activities.

Looking Forward

Whereas the National Council of Teachers of Mathematics's *Curriculum and Evaluation Standards for School Mathematics* (1989) and *Principles and Standards for School Mathematics* (2000) both mentioned the importance of graphing in multiple coordinate systems, as currently written the *CCSSM* (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) only outlines graphing relationships in the CCS. One could argue that the removal of the PCS is the appropriate course of action since students encounter well-documented difficulties with the CCS and instruction on the PCS tends to be shallow and poorly understood by students. On the other hand, not including occasions for working with other coordinate systems, and specifically the PCS, might deprive students of valuable learning opportunities. For instance, having students experience multiple coordinate systems may help students develop a deeper understanding of the CCS, possibly addressing some of the common issues seen in the literature (e.g., viewing graphs solely as pictorial objects to be manipulated). Future research is needed to understand the potential use of multiple coordinate systems to support student learning, and such research should provide a finer-grained characterization of students' mental actions as they construct these coordinate systems and attempt to coordinate their activity among representational systems. The present study illustrates that using multiple coordinate systems can aid pre-service teachers in foregrounding covariational reasoning, and subsequent studies should involve additional pre-service teachers and extend this work to secondary and undergraduate mathematics students.

A limitation of this study is that it included only two students, neither of which encountered significant difficulties over the course of the teaching experiment. Thus, the results speak to the affordances of covariational reasoning when investigating multiple coordinate systesms, but not the development of such reasoning. An area of interest is thus investigating how students with a range of covariational reasoning abiliities work in the PCS, and particularly those students who have not encountered instruction on quantitative and covariational reasoning. Such studies can offer deeper insights into students' *development* of covariational reasoning, including how different images of change might influence such reasoning, in the PCS. For instance, Johnson (2012) has made progress in characterizing nuances in early secondary school students' covariational reasoning. Determining how such nuances evolve over time and gaining insights into the implications of these developments form important areas of study in relation to students' mathematical thinking and learning at the secondary and undergraduate level.

Acknowledgement: Research reported in this paper was supported by National Science Foundation (NSF) Grant No. EHR-0412537 and the University of Georgia (UGA) College of Education Early Career Faculty Grant. Any conclusions or recommendations stated here are those of the authors and do not necessarily reflect official positions of NSF or UGA.

References

Carlson, M. (1998). A cross-sectional investigation of the development of the function concept. In E. Dubinsky, A. H. Shoenfeld & J. J. Kaput (Eds.), *Research in collegiate mathematics education, III. Issues in Mathematics Education* (Vol. 7, pp. 114-162).

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.

Castillo-Garsow, C. W. (2012). Continuous quantitative reasoning. In Mayes, R., Bonillia, R., Hatfield, L. L., and Belbase, S. (Eds.), *Quantitative reasoning and Mathematical modeling: A driver for STEM Integrated Education and Teaching in Context. WISDOMe Monographs*, *Volume 2* (pp. 55-73). Laramie, WY: University of Wyoming Press.

Castillo-Garsow, C. C., Johnson, H. L., & Moore, K. C. (submitted). *Chunky and smooth images of change*.

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

Goldenberg, E.P., Lewis, P., & O'Keefe, J. (1992). Dynamic representation and the development of a process understanding of function. In Harel, G. & Dubinsky, E. (Eds.) *The concept of function: Aspects of epistemology and pedagogy* (pp. 235-260). MAA Notes, Volume 25. Washington, D. C.: Mathematical Association of America.

Johnson, H. L. (2012). Reasoning about variation in the intensity of change in covarying quantities involved in rate of change. *The Journal of Mathematical Behavior, 31*(3), 313-330.

Lobato, J. & Bowers, J. (2000). *Three perspectives in research on functions: Multi-representational, quantitative, and phenomenological.* Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA.

Monk, S. (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* (pp. 175-193). Washington, D.C.: Mathematical Association of America.

Montiel, M., Vidakovic, D., & Kabael, T. (2008). Relationship between students' understanding of functions in Cartesian and polar coordinate systems. *Investigations in Mathematics Learning, 1*(2), 52-70.

Montiel, M., Wilhelmi, M., Vidakovic, D., & Elstak, I. (2009). Using the onto-semiotic approach to identify and analyze mathematical meaning when transiting between different coordinate systems in a multivariate context. *Educational Studies in Mathematics, 72*(2), 139-160.

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

Moore, K. C. (2012). Coherence, quantitative reasoning, and the trigonometry of students. In R. Mayes & L. L. Hatfield (Eds.), *Quantitative Reasoning and Mathematical Modeling: A Driver for STEM Integrated Education and Teaching in Context* (pp. 75-92). Laramie, WY: University of Wyoming.

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

NCTM. (2000). Principles and standards for school mathematics. Reston, VA: NCTM.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Retrieved from http://www.corestandards.org/the-standards/mathematics.

Oehrtman, M. (2008). Layers of abstraction: Theory and design for the instruction of limit concepts. In M. P. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* (pp. 55-64). Washington, D.C.: Mathematical Association of America.

Oehrtman, M., Carlson, M., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. P. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* (pp. 27-42). Washington, D.C.: Mathematical Association of America.

Piaget, J. (1977). *Studies in reflecting abstraction*. Hove, UK: Psychology Press Ltd.

Sayre, E., & Wittman, M. (2007). Intermediate mechanics students' coordinate system choice. *Electronic Proceedings for the Tenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education.* San Diego: San Diego State University.

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.

Thompson, P. W. (1990). A theoretical model of quantity-based reasoning in arithmetic and algebra. Center for Research in Mathematics & Science Education: San Diego State University.

Thompson, P. W. (1994a). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: SUNY Press.

Thompson, P. W. (1994b). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics, 26*(2-3), 229-274.

Thompson, P. W. (1994c). Students, functions, and the undergraduate curriculum. In E. Dubinsky, A. H. Schoenfeld & J. J. Kaput (Eds.), *Research in Collegiate Mathematics Education: Issues in Mathematics Education* (Vol. 4, pp. 21-44). Providence, RI: American Mathematical Society.

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. (2011). Quantitative reasoning and mathematical modeling. In Hatfield, L. L., Chamberlain, S., and Belbase, S. (Eds.), *New perspectives and directions for collaborative research in mathematics education* (pp. 33-56). Laramie, WY: University of Wyoming.

Thompson, P. W. (in press). In the absence of meaning. In K. Leatham (Ed.), *Vital directions for research in mathematics education*. New York: Springer.

STUDENTS' EMERGING UNDERSTANDINGS OF THE POLAR COORDINATE SYSTEM

Teo Paoletti
University of Georgia
paolett2@uga.edu

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

Jackie Gammaro
University of Georgia
jgammaro@uga.edu

Stacy Musgrave
University of Georgia
smusgrav@uga.edu

The polar coordinate system (PCS) arises in a multitude of contexts in undergraduate mathematics. Yet, there is a limited body of research investigating students' understandings of the PCS. In this report, we discuss findings from a teaching experiment that explored pre-service teachers' developing understandings of the PCS. We specifically identify several issues that arose spontaneously as we worked with the students. For instance, we illustrate how students' angle measure meanings influenced their construction of the PCS. We also discuss ways in which the students' understandings of the Cartesian coordinate system (CCS) became problematic as they transitioned to the PCS. As an example, mathematical differences between the polar pole and Cartesian origin perturbed some students. Collectively, our findings highlight the emergent nature of students' coordinate systems.

Key Words: Polar coordinates, Coordinate systems, Pre-service secondary teachers, Teaching experiment, Quantitative reasoning

Introduction

In addition to having many mathematical applications, polar coordinates have a multitude of real-world uses, several of which are found in physics and engineering (Montiel, Wilhelmi, Vidakovic, & Elstak, 2009; Sayre & Wittman, 2007). Often, students are first introduced to polar coordinates in pre-calculus, where the emphasis tends to be on conversions between Cartesian and polar coordinates and recognizing 'special' graphs (e.g., limacons and cardioids) (Montiel, Vidakovic, & Kabael, 2008; Montiel et al., 2009). Polar coordinates later arise in single-variable calculus where students are asked to determine areas contained by polar curves using integration. Following the use of polar coordinates in single-variable calculus, multi-variable calculus extends polar coordinates to spherical and cylindrical coordinates in order to model three-dimensional objects and situations.

Available research on students' understandings of polar coordinates has emphasized problematic meanings that students have for the *polar coordinate system* (*PCS*). For instance, Montiel and colleagues (2008) identified that students' understandings of the PCS are often tied to their activity within the *Cartesian coordinate system* (*CCS*) (e.g., students apply the vertical line test in the PCS). Such observations generate a need for better understanding students' *construction* of the PCS. Although the PCS and CCS are both coordinate systems used to represent relationships, these systems have different conventions that students must construct and coordinate if they are to productively use the PCS.

In order to better understand students' construction of the PCS, we conducted a teaching experiment (Steffe & Thompson, 2000) with a group of undergraduate students enrolled in a course for pre-service secondary mathematics teachers. The following research questions drove the study:

(a) What ways of thinking do students engage in when constructing the PCS?
(b) What issues arise during students' construction of the PCS?
(c) How do students utilize the CCS while constructing the PCS?

In the present work, we discuss issues that arose as our participants constructed the PCS and the ways of thinking that led to these issues. For example, students' meanings for angle measure influenced their conception of the PCS. Further, and consistent with previous research, we found that some students relied on rules and conventions from the CCS when plotting points or interpreting graphs in the PCS. Against the backdrop of such issues, we highlight the emergent nature of students' PCS.

Background

Available research, which is scarce, has shown that collegiate students often struggle with the PCS, both in mathematical and science/engineering problems (Montiel et al., 2008; Montiel et al., 2009; Sayre & Wittman, 2007). Sayre and Wittman (2007) illustrated that engineering students leverage the PCS at varying degrees when working in contexts relevant to the PCS. The authors concluded that students' understandings of the PCS influenced their tendency to simplify real-world problems using the PCS. Namely, those students with weaker understandings of the PCS relied on the CCS, often to their own detriment, in situations more suitable for the PCS. Such findings point to the importance of students constructing the PCS in ways that support their ability to interpret and solve problems that arise in real-world contexts, particularly in engineering and applied fields.

Montiel et al. (2008) investigated second semester calculus students' understandings of the PCS in relation to the students' function concept. They illustrated that many students' function concept relied on the vertical line test, which became problematic in the PCS. For instance, students applied the vertical line test to the polar graph of $r = 2$ (e.g., a circle) to conclude the graph was not a function. Others first converted the equation to its CCS form (e.g., $x^2 + y^2 = 4$) and then applied the vertical line test to conclude that the graph does not represent a function. Collectively, Montiel and colleagues' findings highlight students' meanings for concepts like function can be tied to a particular coordinate system, which, in turn, can influence their construction and use (e.g. representing covarying quantities) of other coordinate systems.

Although previous research provides several important findings, a limitation is that these studies have only examined student understanding *after* instruction had taken place. Because of this, each study utilizes questions that presume an understanding of the PCS (e.g., asking students if a polar curve or equation defines a function). In the present study, we take a fundamentally different approach and explore students' construction of the PCS during activities designed to help them to continually develop and reconstruct meanings for a coordinate system.

In their work on students' construction of the CCS, Piaget Inhelder, and Szeminska (1960) described children using ideas similar to those required by the PCS. In the study, the researchers tasked children with duplicating the exact location of a point on a piece of paper to another piece of paper across a table, so that when the two pieces of paper were overlaid, the points would overlap exactly. Many children found the distance from a corner of the paper to the point and attempted to use only this distance to mark the exact location of the point on the second piece of paper. Some of these children also tried to maintain the inclination of their ruler while preserving the distance from the corner of the paper. These students attempted this by "moving the ruler in short steps, keeping the inclination uniform throughout" (Piaget et al., 1960, p. 163). Such activity is consistent with using a directed length combined with an angle of constant openness to find the location of a point in space, thus creating an intuitive form of the PCS.

Given their focus on the CCS, Piaget et al. (1960) did not give extensive consideration to the students' intuitive notions of the PCS. And although the present study does not address young children's construction of the PCS, we do explore students' formalizing two quantities (e.g., a directed length and the openness of an angle) into a coordinated system for describing points in a plane. The results of Piaget et al. (1960) indicate that the PCS may be an intuitive system at some level, but we conjectured that particular conventions of the PCS could prove to be emergent problems for students as they construct the PCS and attempt to draw on their CCS conceptions. For instance, whereas the CCS is based on two directed lengths, the PCS includes a coordinate that represents a dimensionless quantity–angle measure–and thus how a student conceives the PCS will be reliant on their ability to coordinate a dimensionless quantity with a directed length.

Quantitative Reasoning and the Polar Coordinate System

In addition to leveraging available research on the PCS, we also draw on theories of *quantitative reasoning* (Smith III & Thompson, 2008; Thompson, 2011). Quantitative reasoning takes the stance that quantities are cognitive constructions and therefore should not be taken as a given (Thompson, 2011). Previous research has indicated that quantitative reasoning is crucial to students' understandings of numerous topics in undergraduate mathematics including rate of change (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002), function (Oehrtman, Carlson, & Thompson, 2008), trigonometric relationships (Moore, 2012), exponential relationships (Castillo-Garsow, 2012; Confrey & Smith, 1995), and the fundamental theorem of calculus (Thompson, 1994). Specific to coordinate systems, quantitative reasoning provides a lens by which to make sense of students organizing a coordinate system in terms of a systematic coordination of two quantities (e.g., two direct lengths for the CCS, one directed length and an angle measure for the PCS).

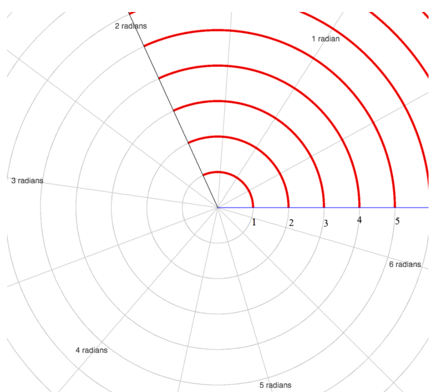


Figure 1. Polar coordinates and the angular component defining an equivalence class of arcs.

As an illustration of quantitative reasoning in the context of the PCS, consider the set of points defined by $(r,2)$. The radial component defines a directed length r units from the pole and the angular component defines a set of points forming a ray at an angle of measure 2 radians counter-clockwise from the polar axis. When viewing the set of points defined by $(r,2), r>0$, and with an understanding of angle measure rooted in arcs, one can conceive *every* arc subtended by the formed angle as having a length of 2 radii when measured in a magnitude equivalent to that circle's radius (Figure 1); the ray defines an equivalence class of arcs upon which the coordinate system is based. Moore (2012) illustrated that quantitative reasoning plays a central role in students' meanings for angle measure, including their ability to quantify angle measure such that a measure entails an equivalence class of arcs. As we later demonstrate, without understanding the ray as defining an equivalence class of arcs tied to radii measures, a student can be left with arcs of different lengths being used to represent what should be a constant value. This image can lead to interpreting points along a ray as having different angular (arc) values.

Methods and Subjects

Our intention in the study, based on a radical constructivist epistemology (Glasersfeld, 1995), was to build and test models of the students' thinking. To accomplish this goal we conducted a teaching experiment (Steffe & Thompson, 2000). The context for the teaching experiment was a mathematics education content course for pre-service secondary teachers at a large state university in the southeast United States. Theories of quantitative reasoning (Smith III & Thompson, 2008; Thompson 1990) and covariational reasoning (Carlson et al., 2002) informed the course. Prior to the teaching experiment, the students in this study were exposed to a treatment of angle measure and trigonometric functions described by Moore (in press, 2012). The students in the class were third-year (in credits taken) undergraduates who had taken Calculus II and at least two other mathematics courses. Time in class was mostly spent with students working in groups, with the instructor then coordinating whole class discussions based on students' presentation of their work.

In addition to capturing whole class discussions (21 total students), we focused our data collection on four students as they worked in pairs. These four participants (Katie, Jenna, John, and Steve) were chosen from a pool who volunteered. The students were paired based on the results from a research-based pre-assessment (Carlson, Oehrtman, & Engelke, 2010) given to students at the beginning of the content course and homework evaluations for the first month of this class. Based on this data, we paired two higher performing students (Katie and John) and two lower performing students (Steve and Jenna). We conjectured pairing in this manner would help us identify different ways students construct the PCS.

Each pair of students worked with one teacher-researcher while an observer-researcher videotaped their interactions to capture their writing, actions, and words. During whole class instruction, the observer-researchers videotaped all interactions and conversations that occurred, including those that the pairs of students might have as asides during the whole class discussions. Over the course of three weeks, the teaching experiment spanned five class sessions, each lasting approximately one hour and fifteen minutes.

All of the videos were transcribed and then reviewed by the research team. When analyzing the data, we used a combination of an open and axial coding approach (Strauss & Corbin, 1998) and conceptual analysis (Thompson, 2000). We first analyzed each pair of students' words and actions in order to characterize their thinking. Coupled with the classroom conversations, we further examined common and often unexpected themes which arose in our analyses. Upon our characterization of these themes, we searched the data for instances confirming or contradicting our models of the students' thinking. We sought to build more viable models (Steffe & Thompson, 2000) of the students' thinking, including shifts in their thinking, through this iterative process.

Results

While all students in the class had taken mathematic courses beyond Calculus II and acknowledged some familiarity with the PCS, their activity during the teaching experiment suggested that constructing the PCS and comparing various properties of the system to properties of the CCS were novel experiences. In this section, we illustrate specific instances in which students faced problematic situations when constructing and working with the PCS. Due to the focus on problematic situations, some of which persisted over multiple teaching sessions, we do not present the data in chronological order. We provide background information relative to the place in the teaching experiment when necessary for clarification.

Through our conceptual analysis of the data, we noted various themes in the students' activity around which we organize this section. The first theme centers on the importance of students' radian angle measure meanings while constructing the PCS. The second theme deals with characteristics of the PCS and CCS, specifically how the students handled various conventions of the systems.

Constructing the PCS and the Importance of Radian Angle Measure

The first activity in the teaching experiment asked the students to determine what information a surface radar system on a ship, with the ship defining a center point, might give in order to determine the exact location of another object at sea (Figure 2). We posed that the system provides how far a detected object is from the ship and asked the students to determine a second quantity that would define an exact location of the object. Our goal was for students to realize a distance from the fixed point and an angle measure with a defined reference ray and rotation direction suffices to provide an exact location in the plane. At the point in which the students determined these quantities, we intended to task them with constructing a coordinate system based on the quantities (e.g., the PCS).

Consider a boat at sea using sonar to determine the location of other objects.

1) If the sonar device detects an object that is 2 miles away, draw all locations on the screen that could correspond to the object.
2) If the sonar device detects an object that is 3.5 miles away, draw all locations on the screen that could correspond to the object.
3) What other measurement might you add to the device that will allow the crew to convey the exact locations of the objects from parts (a) and (b)? Illustrate multiple locations for an object and the corresponding values for this location.
4) Draw a coordinate system that conveys how far any object is from the boat and the measurement identified in part (c). Identify and label 3 points on your coordinate system.

Figure 2. Sonar problem presented to the students.

I. The Importance of Radian Angle Measure

When comparing the two pairs of students' construction of the PCS over the course of the task, it became apparent that their angle measure meanings influenced their work on the task. As mentioned above, the students had received instruction on radian measure based in research on the topic (Moore, 2012, in press). Unsurprisingly, the students' work on the sonar problem established that the students had developed different understandings during the angle measure lessons. Namely, Steve and Jenna's activity suggested that they had not constructed meanings compatible with those that we had aimed to develop.

Steve and Jenna experienced difficulty coordinating the angular component in the PCS, particularly when considering different arc lengths as corresponding to the same angle measure. After choosing angle measure as a second quantity for the radar system, the students described the process of determining the radial distance-angle measure pair for a point in the plane (Excerpt 1).

Excerpt 1

Jenna and Steve discuss how to label a point on a circle using angle measure

(*The students have circles with radii of 2 miles and 3.5 miles and, to the researcher, rays defining 1 radian and 2 radians on their board–see Figure 3*)

Jenna: I was saying how many of our radians can we fit (*tracing along an arc of the 3.5-mile radius circle from the line due east to a line due north*).

Steve: That's what I was going to say. If, if you made it into…

Jenna: (*using a piece of waxed string that was used as the radius of the 2-mile circle to measure the arc length from the line due east to the line due north on the 3.5-mile radius circle)* One… two…. Oh, that's close to three.

Steve: Yeah.

Jenna: So we know…

Int.: So you're saying if we were, you're focusing if we were right here (*indicating the location where the 3.5-mile radius circle intersects the line due north*)?

Jenna: Yeah, well that's, I'm just saying that's our ninety degrees.

Int.: Oh, okay. So ninety degrees.

Jenna: So if we have three six nine twelve (*pointing to locations on the 3.5-mile radius circle corresponding to 90, 180, 270, and 360 degrees*), we have twelve… rad… radians, right?

(*After working on some calculations the conversation regarding radians continues*).

Int.: So how are you using radian right now? The word radian, how are you using that?

Steve: We're using radian in the fact that this (*pointing to a length equal to the radius of the 3.5-mile circle*), were making the three and a half miles into one unit. And then, we're, this is one (*referring to the radius length of 3.5 miles*). And then we're taking that one unit…Whatever, oh crap. We didn't, we used this (*the length of 2 miles*) instead of …

Jenna: Two miles.

Steve: Instead of three and a half. Well you're right, this (*the 2 mile length*) would have worked for this (*referring to the 2-mile radius circle*) though, because say you would've went like this (*using a waxed string with a length of 2 miles to wrap around the arc of the 2-mile radius circle*), and then you would've kept going.

Jenna: What do you mean, no we didn't? Because this angle measure (*referring to the angle labeled theta in Figure 3*) whether it's here (*pointing to the arc intersected by the angle on the 2-mile circle*) or here (*pointing to the arc intersected by the angle on the 3.5-mile circle*), is the same angle. This is, our one radian is two miles (*showing the waxed string has a length of 2 miles*).

Steve: Okay. That's right. Alright, you use this (*referring to the 2-mile radius length*) regardless. You're right, that makes sense. Okay.

Int.: So what are you saying now? You use that (*inaudible*)

Jenna: So we're saying, but we're going to have to multiply whatever it is. Like if we wanted to know our arc length (*tracing along the subtended arc on the 3.5-mile circle*), we'll have to multiply by two in the end. Because we use this (*the 2-mile length*) as one radian and one radian is equal to two miles.

In this interaction, Jenna conceives a "radian" as a fixed magnitude, which in this case is 2 miles (lines 30-41). For instance, when measuring the radian measure of an arc on the circle of 3.5 miles, she used a magnitude equivalent to 2 miles (the radius of the first drawn circle) to conclude that a 90-degree angle corresponds to approximately 3 radians. From there, she suggested that there are 12 radians in the full circle (line 15-17). Although Steve did suggest that they need to coordinate their unit magnitude with the radius of the circle (lines 20-24), which is a foundational way of thinking for radian measures as equivalence classes (Moore, 2012, in press), he was quickly swayed (lines 35-36) by Jenna's claim that "one radian" is fixed at two miles.

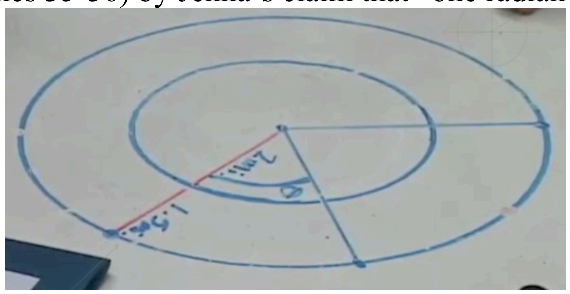


Figure 3. Steve and Jenna's circles with radii 2 and 3.5 miles

By choosing a fixed magnitude for measuring arcs, Steve and Jenna created a paradox in their emerging coordinate system. Specifically, this outcome led to the students interpreting arcs on different circles but subtended by the same angle as producing different coordinate measures. To illustrate, Figure 4 represents an analogous situation where the radial unit magnitude is also used as the magnitude for measuring arcs, regardless of circle. Specifically for the angle shown in Figure 3, the coordinate pair (1, 1) represents the point where the terminal ray intersects the circle with a radius of 1 unit, while the coordinate pair (2, 2) represents the point where the terminal ray intersects the circle with a radius of 2 units. This creates an issue as an angle of constant openness has both a measure of 1 radian and 2 radians (and, in fact, an infinite number of radian measures). Because of such an outcome, the pair questioned using angle measure, as they had conceived, as an appropriate quantity for conveying the location of a point in the plane.

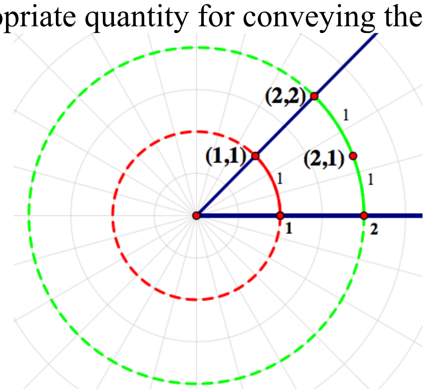


Figure 4. A representation of using a fixed magnitude to measure arcs.

As the teaching experiment progressed, we worked with the students to develop radian measures as defining an equivalence class of arcs. As the pair constructed such an understanding, they both made strides in their PCS construction. However, this progress was neither trivial nor quick. For instance, later in the first teaching episode Jenna asked, "I get confused when we were saying this (*indicating the arc length of 3.5 miles on the 3.5-mile radius circle*) was a radii, but then that's also an angle measure. It's an arc length and an angle measure?" This quote illustrates Jenna's continuing struggle to coordinate the relationship between the radius of a circle, subtended arc lengths, and angle measures. During the second teaching episode, Jenna added, "See that's what confuses me, because how can we say that the length, I mean it's, I understand that [the arc length] is the same length as one radius, one radian. But the angle measure is also one radian, so they're all (*referring to several subtended arcs*) one radian?" After a discussion that entailed identifying that "one radian" refers to *all* of the arcs subtended by the angle, Jenna explained that she had difficulty reasoning about an angle measure in terms of measuring arcs due to having to coordinate a unit magnitude that depends on that particular circle.

Later in the second session, and after the above quote, Jenna continued to fix a radian as a fixed magnitude, yet was aware of the paradoxes this created. After the pair constructed circles with radii measuring 1, 2, 4, and 5 inches, they drew rays corresponding to rotations of 1, 2, 3, and 4 radians as shown in Figure 5. Jenna then fixed the "radian" length to one inch, leading her to question how it is possible that an angle measure can be one inch for all circles (Excerpt 2).

Figure 5. Steve and Jenna's circles with radii 1, 2, 4, and 5 inches.

Excerpt 2

A discussion of what it means to measure one radian

Jenna: One inch (*tracing the arc length on the one-inch circle intersected by the one radian ray*), that means we have, this arc length is one inch *(pointing to the traced arc length)*, that's one inch (*pointing to the radius*), and then the angle measure (*motioning over the angle created near the pole*) is one inch…
Steve: One radian, but one radian would be one inch.
Jenna: One inch. You can have a one-inch angle measurement?
(*Jenna continues to state her dissatisfaction an angle being measured in inches*)
Int.: Okay so what is it about one inch that would pose a problem, why can't we say an angle has a measure of one inch?
Jenna: Cause where do you measure the one inch?
Int.: So say more, what do you mean by, what do you…
Jenna: You can't just set like. Angle measures have like… cause it's not going to be one inch all the way, like as the angle measure goes out (*moving fingers apart as she traces them along the two highlighted rays*).
Steve: Oh yeah, it's the openness of the angle.
Int.: Okay, cause what happens when we go… if we say, we say an angle has a measure of one inch what happens when we go to this circle (*referring to the circle with a radius of 2 inches*)?
Jenna: It (*referring to the arc length)* still has, it should still have one, the measure of one inch, but that's not (*identifying that the arc length is not 1 inch on the circle with a 2 inch radius*)…
Steve: One inch, because the one-inch is corresponding to this arc length (*referring to the arc length on the circle with a 1 inch radius*).
Int.: And so what would one inch be on this [circle] (*referring to the circle with radius 2 inches*)?
Steve: It'd be half.
Jenna: Half.
Int.: It would be half, right. So we'd get an angle of different measures.

29 Steve: So that's why you said it's a one unit.
30 Int.: That's why we want…
31 Jenna: When you talk of angle measures you have to say one radian.
32 Int.: Exactly.
33 Jenna: You can't just, you can't convert it to inches.

The above interaction exhibits an indicator of Jenna reconciling that a fixed unit length will not be sufficient to measure along arcs on different circles, as this will result in different measures for the same angle (lines 19-21, 31-33). After this interaction, the interviewer asked the pair about the meaning of the quantity given by the quotient of a subtended arc length and the radius of the circle on which the arc length lies. Jenna's words and actions indicated a developing understanding that for a given angle measure this quotient would be equal for every circle centered at the vertex of the angle. At this point far into the second session, we see instances where Jenna was beginning to recognize the need of changing the unit magnitude when changing circles. Collectively, the students' difficulties with angle measure corroborate Moore's findings (2012) about students' images of angle measure being problematic, while highlighting the importance of students' meanings for angle measure when constructing the PCS.

PCS and CCS Characteristics, Conventions, and Influences

While we conjectured that a robust understanding of radian angle measure was a prerequisite for developing flexible PCS meanings, other themes that emerged as students attempted to connect their CCS meanings to the PCS were unexpected. As one example, issues arose due to differences among the origin of the CCS and the pole in the PCS. As another example, the students faced perturbations with various conventions related to function and coordinate pair ordering.

I. The origin vs. the pole

One theme that emerged during the teaching experiment was the students' evolving conceptions of the pole in the PCS. The pole (or "origin" as referred to by the students) became problematic when we provided a graph (Figure 6) and asked them to determine an equation to define the graph (e.g., $r(\theta)=2\theta-0.5$, $\theta\geq 0$). After finding a constant rate of change for the relation, Steve used the fact that the graph passed through the "origin" to determine $r(\theta)=2\theta+0$, a fact he recalled from the CCS (e.g., a graph intersecting the "origin" corresponds to an intercept of 0). However, when the pair tested their rule with another point, they found that the parameter was -0.5 rather than zero (e.g., $r(\theta)=2\theta-0.5$). This led the pair to state that the given graph was misleading in that it did not pass through the "origin," but instead passed through a point very close to the origin.

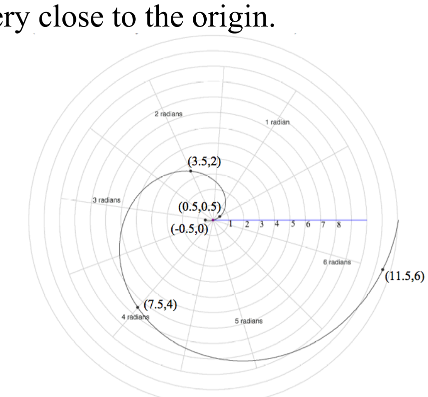


Figure 6. The graph presented to the students.

This assumption satisfied Steve and Jenna, but perturbed us and we decided to raise the issue of the origin in the whole class discussion. Taking a quick survey of the class, a majority of students believed the graph did not pass through the origin. Steve was the first to make an argument stating, "It doesn't go through the origin. 'Cause that wouldn't match our function (*referring to the equation*)." When probed to justify his thoughts, Steve argued, "If you use the origin zero-zero, then if your radius is zero and your theta is zero, there's no way you can get zero unless b is zero." Other students in the class began to vocalize their agreement that the graph did not pass through the origin (Excerpt 3).

Excerpt 3

Classroom discussion if the function $r(\theta) = 2\theta - 0.5$ *included the origin in the PCS.*

Student A:	If you plug zero for theta, your r of theta will not be zero.
Int.:	Okay so you plug in zero for theta you're not going to get that (*referring to the point (0, 0)*).
Student B:	Okay, for, it wouldn't be a function if you have two solutions, like two outcomes, when you enter zero. Like, the radians, if you enter theta at zero then you get two different radii (*referring to the points (-0.5, 0) and (0, 0)*) then it's not a function.
Int.:	So you're saying we'd have another point associated with zero if it did pass through the origin.
Student B:	Yeah, and if this is, we're graphing a function. It wouldn't be a function.
Int.:	Okay.
(*A third student makes an argument against the graph going through the origin based on the constant rate of change of the function*)	
John:	Can I say why it does go through the origin?
Int.:	You can say why it does, why you think it goes through the origin.
John:	Any time the radius is equal to zero, the swirl goes through the origin, doesn't matter theta.
Katie:	Because there is no angle if the radius is zero.
Int.:	What do you guys think of that?
(*Student's making noises to indicate that they agree with John and Katie's statements*)	
Steve:	Ohh… way to throw a wrinkle in it
Int.:	So what do you guys think? So John, say that a little louder, maybe, can… say that a little louder John.
Steve:	(*an aside to Jenna*) So that means our function is wrong?
John:	If at any point on r theta, um…
Int.:	So you're saying if we have any point r theta and we do what…
John:	Where r is equal to zero, it doesn't matter what theta is, because if you don't have a radius away from the origin, you can't have an angle anyway.
Int.:	So what if we have any point that's zero, and then I give you any angle measure, where are we graphing that point?
John:	(*in unison with many other students*) Through the origin.

Like Steve and Jenna, a majority of the students believed that passing through the "origin" in the PCS required that the function include the point (0, 0) (lines 1, 4-7). Thus, obtaining a rule that did not entail this pair led the students to conclude that the given graph could not pass through the "origin" as they had conceived it. As part of their argument, the students used the function notion that each input only has one associated output to argue (correctly) that it was impossible for the points (0, 0), and (-0.5, 0) to both be coordinate pairs associated with the function. The students' arguments are not incorrect from a function standpoint, but instead the problem rests within their conception of the PCS "origin"; by extending the coordinate pair for the origin in the CCS to the PCS, the students created a problematic situation.

While most students eventually agreed with John's reasoning, Steve remained perturbed, asking Jenna about the correctness of their function (line 25). He later stated, "It's so hard to grasp that it goes through the origin but it's not zero-zero. I feel like that's ingrained in our mind." John also continued to focus on the PCS pole as the class continued. After the above conversation, John spontaneously asked the teacher-researcher, "Would *r* of theta equals theta squared plus one not go through the origin?" John's question indicates that he developed a self-directed goal, namely to think of a function that would not pass through the pole in the PCS. In all, the classroom conversation regarding the pole not only led to the students recalling their function meanings but also their understandings of the underlying structure of the CCS and PCS, including relationships between the origin in the CCS and the pole in the PCS.

II. Conventions

The pole is only one example where the students' developing PCS understandings were influenced by their previous experience with the CCS and the conventions that differed across the coordinate systems. When first graphing functions in the PCS, students struggled with the convention of a coordinate point (r,θ) representing (output, input) when working with functions of the form $r(\theta)$. This convention created issues for students when they attempted to make sense of other ideas, like rate of change. For instance, while trying to find the rate of change of the relationship in Figure 6, Steve and Jenna initially relied on memorized formulas (Excerpt 4).

Excerpt 4

Steve and Jenna attempt to use a formula to find the rate of change of the function

Steve:	Yeah, we can use, what, we can use *y-two* minus *y-one*. Find the slope, since it's the same as using the Cartesian.
Jenna:	Sure, sure.
Steve:	Okay.
Jenna:	So well do like, um, change in, change in *r* over change in theta (*writing* $\Delta r/\Delta\theta$) or is it vice versa?
Steve:	Well see, theta is in the *y*, *r* theta (*referring to how the coordinate point is written*). Is it, is it…
Jenna:	Yeah, it's *r* theta.
Steve:	Yeah. So changes in *y* over changes in *x*, is that what it is? (*Jenna changes the formula to be* $\Delta\theta/\Delta r$)
Jenna:	Yeah.
Steve:	Yeah, there we go. Awesome. This… that seems right? Okay.

In this interaction, we see the pair relying on the CCS convention of (input, output) to find the rate of change of the function, rather than focusing on how the quantities were being treated relative to the function (*r* as a function of *θ*). Shortly after this, Jenna transitioned to thinking about rate of change in terms of input and output quantities.

Excerpt 5

Jenna relying on quantities to find the equation of a line

Jenna:	You said change of output over change in input. This is output (*writing "output"*

over the r in her polar ordered-pair (r,θ)).
Steve: You're right.
Jenna: And this is input (*writing "input" over the θ in her polar ordered-pair (r,θ)*).
Steve: Yep.
Jenna: So it would be change in *r* over change in theta (*rewriting* Δ*r*/Δ*θ*).
Steve: That's right cause we did it a different way. You could make it theta *r* (*referring to the ordering of the coordinate pair*) if you wanted to, but we made it *r* theta.

In the above example, it was critical for Jenna and Steve to be concerned with the relationship between quantities rather than the rules they memorized in the CCS. Initially the pair attempted to find the rate of change of the function by applying the slope formula of $(y_2 - y_1)/(x_2 - x_1)$ and then replacing the *y*'s with the corresponding second polar coordinate, θ, and the *x*'s with the corresponding first polar coordinate, *r*. It was not until Jenna recalled the relationship between quantities, namely rate of change as changes in output over changes in input, that the pair rewrote the formula to represent this relationship. Interestingly, Steve explained that they could have switched the order of their coordinate pairs so the familiar conventions from the CCS would have applied (lines 7-8). They opted for the order established previously, with Steve highlighting the arbitrary nature of this decision: "I dunno why we did, but we did". Immediately after this conversation, Jenna stated that she was "really confused on the homework" due to treating the first coordinate as the output of a function in the form $r(\theta)$.

In hindsight, the students' discomfort with the polar convention of (r,θ) as (output, input) is not all that surprising. In fact, Montiel et al. (2008) noted a similar outcome in their study where students had trouble identifying the independent variable when deciding if a graph represented a function. Conventions relative to input and output variables also emerged when equations were later given with θ as a function of *r*. For instance, when asked to graph $\theta = r^2$, Steve and Jenna chose values of θ and found corresponding values of *r* such that $r = \theta^2$. As the students progressed on the task, Steve eventually claimed, "It's theta equals *r* squared, so wouldn't *r* be the square root of theta, right?" Steve then rewrote the equation with θ as the input of the function (e.g., $r(\theta) = \sqrt{\theta}$). After substituting in the point (4,2) to $\theta = r^2$ to convince Jenna of their error, Jenna claimed, "We just did the math backwards."

Excerpt 6

Steve and Jenna's conversation about polar conventions

Jenna: If we did that backwards, erase it.
Steve: Yeah, let's see. Cause he switched the inputs and outputs on us. So that means we have to figure out, we have to do a table, like…
Jenna: Uh, why don't we just do it backwards, why don't we do
Steve: So *r* equals square root of theta (*attempting to find values of* $r = \sqrt{\theta}$).
Jenna: Why don't we just do, put in for *r* and get theta.
Steve: Oh yeah, that's a, a perfect idea, that's fine, that works.

After stating that the roles of the variables had switched in the given rule (e.g., the given rule implied that θ was the output and *r* the input), which conflicted with their prior experiences that defined (r,θ) as (output, input), Steve's desire was to rewrite the equation in a manner suitable for this convention. Conversely, Jenna decided to switch the roles of *r* and θ. She was comfortable ignoring their previously defined convention and worked fluidly with choosing either quantity as the input. After Jenna suggested this, Steve identified such conventional practices as being arbitrary decisions, but his later actions indicated that his preference was on the convention (e.g., (r,θ) as (output, input)) he adopted from earlier experiences in the PCS.

On this same problem, John and Katie also wanted to rewrite the equation as *r* in terms of θ.

Excerpt 7

Katie and John attempting to graph $\theta = r^2$

John: So the Cartesian one is gonna… be *y* equals square root of *x* right?
Katie: Yea.
(*Katie draws a Cartesian plane and the equation y equals square root x*)
Katie: So we needed, how does…
John: Or should we do it, not necessarily like a function, should it be positive and negative (*inaudible*).
Int.: Whatever you guys think, approach it however you want.
Katie: So we need to do… Hold on, what was the function again?
Int.: Theta equals *r*-squared
(*Both write the equation on the whiteboard*)
Katie: So let's just plug in, it's easier just to plug in *r*'s.
John: Um but then… We'd have to work back to get the theta.
(*Katie creates a t-table with theta as the input and begins plugging in values.*)
John: Sure. Wait, ahh… is that right? Why was I thinking it was square root of two*? r* equals square root of theta. Yeah okay, your right sorry…

Similar to Steve and Jenna's actions, John and Katie attempted to modify the given function so that it was in the form $r(\theta)$. However, they quickly reconciled this issue and identified that this was not a necessary action.

Conclusions and Implications

Researchers have suggested that students have superficial meanings of the PCS after their precalculus and single-variable calculus level experiences (Montiel et al., 2008; Montiel et al., 2009). Our current research uncovers some possible causes of these superficial understandings. One potential source is the underlying structure of the PCS. In the CCS, directed lengths form both coordinate quantities, while in the PCS an angular component forms one of the quantities. It follows that students' angle measure meanings, including conceiving radian measures as an equivalence class of arcs, emerged as central to the students' construction of the PCS. Jenna provides an example of the struggles a student encounters when attempting to develop a coordinate system that relies on radian measure without having radian measure meanings that support his or her construction of the PCS. This and previous research (Moore 2012) indicates the importance of pre-calculus and calculus teachers not taking their students' understandings of radian measure as a given when developing the PCS.

In addition to issues with angle measure, our results support previous findings concerning the influence the CCS has on students' PCS understandings. Previous research (Montiel et al., 2008; Montiel et al., 2009) has identified students' (mis-)use of the vertical line test in the context of the PCS. Our results indicate that other issues stem from certain features of the PCS. For instance, it was important that the students understand the pole of the PCS as represented by an infinite number of coordinate pairs. In the case that the students conceived the pole as represented by a unique pair (e.g., (0, 0), like the CCS origin), they had difficulty reconciling a function's rule and PCS graph.

Underlying our results is the interplay between students' PCS and CCS meanings. Although we emphasize the importance of not taking radian angle measure as a given, our results also indicate that no coordinate system should be taken as a given when working with students. In the present study, the students' construction of the PCS was an emergent process that occurred when working on different problems that asked them to do various activities in the PCS. Further, we noticed by having students work in the PCS, they often revisited their meanings for the CCS. For instance, during the course of the teaching experiment many students (e.g., Steve) realized that the origin as (0, 0) was a specific property of the CCS, not an absolute rule for all coordinate systems. Further, our results indicate potential issues that can arise when students only experience one coordinate system. When attempting to find the rate of change of a graph, Steve and Jenna initially use a formula based on CCS conventions, as this formula had proved viable for them in their previous experiences. By introducing the PCS, the pair's attention was drawn to interpreting their solution in terms of a relationship between quantities in order to derive a formula that was consistent for their rate of change meanings.

Our findings, in combination with previous research (Montiel et al., 2008; Montiel et al., 2009), suggest that if students' experiences are predominantly within one coordinate system, then their understandings for related concepts (e.g., function and origin) can become inherently tied to conventions of that coordinate system (e.g., function means execute vertical line test). Our study also illustrates that introducing a new coordinate system can help perturb such meanings as students attempt to reconcile similarities and differences between coordinate systems. Future researchers may be interested in exploring how to use multiple coordinate systems in order to improve students' representational activity and meanings. For instance, our research suggests that using multiple coordinate systems can foreground a focus on the quantities that are foundational to the coordinate systems. Future research that investigates using a multitude of coordinate systems in the teaching of mathematics might prove useful for determining critical ways of reasoning about function and quantitative relationships including how to promote these ways of reasoning.

Acknowledgement: Research reported in this paper was supported by National Science Foundation (NSF) Grant No. EHR-0412537 and the University of Georgia (UGA) College of Education Early Career Faculty Grant. Any conclusions or recommendations stated here are those of the authors and do not necessarily reflect official positions of NSF or UGA.

References

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.

Carlson, M., Oehrtman, M., & Engelke, N. (2010). The precalculus concept assessment (PCA) Instrument: A tool for assessing reasoning patterns, understandings, and knowledge of precalculus level students. *Cognition and Instruction, 28*(2).

Castillo-Garsow, C. W. (2012). Continuous quantitative reasoning. In Mayes, R., Bonillia, R., Hatfield, L. L., and Belbase, S. (Eds.), *Quantitative reasoning and Mathematical modeling: A driver for STEM Integrated Education and Teaching in Context. WISDOMe Monographs*, *Volume 2* (pp. 55-73). Laramie, WY: University of Wyoming Press.

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

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

Montiel, M., Vidakovic, D., & Kabael, T. (2008). Relationship between students' understanding of functions in cartesian and polar coordinate systems. *Investigations in Mathematics Learning, 1*(2), 52–70.

Montiel, M., Wilhelmi, M., Vidakovic, D., & Elstak, I. (2009). Using the onto-semiotic approach to identify and analyze mathematical meaning when transiting between different coordinate systems in a multivariate context. *Educational Studies in Mathematics, 72*(2), 139–160.

Moore, K. C. (2012). Coherence, quantitative reasoning, and the trigonometry of students. In R. Mayes, & L. L. Hatfield (Eds.), *Quantitative reasoning and mathematical modeling: A driver for STEM integrated education and teaching in context* (pp. 75–92). Laramie, WY: University of Wyoming.

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

Oehrtman, M., Carlson, M., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. P. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Teaching in Undergraduate Mathematics Education* (pp. 27-42). Washington, D.C.: Mathematical Association of America.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. Oxford England: Basic Books.

Sayre, E., & Wittman, M. (2007). Intermediate mechanics students' coordinate system choice. *Electronic Proceedings for the Tenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education,* San Diego.

Smith III, J., & Thompson, P. W. (2008). Quantitative reasoning and the development of algebraic reasoning. In J. J. Kaput, Carraher D.W. & Blanton M.L. (Eds.), *Algebra in the early grades* (pp. 95–132). New York, NY: Lawrence Erlbaum Associates.

Steffe, L., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh, & Kelly A.E. (Eds.), *Research design in mathematics and science education* (pp. 267–307). Hillsday, NJ: Erlbaum.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.)*. Thousand Oaks, CA US: Sage Publications, Inc.

Thompson, P. W. (1990). A theoretical model of quantity-based reasoning in arithmetic and algebraic. Center for Research in Mathematics & Science Education: San Diego State University.

Thompson, P. W. (1994). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics, 26*(2-3), 229-274.

Thompson, P. W. (2000). Radical constructivism: Reflections and directions. In L. Steffe, & P. W. Thompson (Eds.), *Radical constructivism in action: Building on the pioneering work of Ernst von Glasersfeld* (pp. 412–448). London: Falmer Press.

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.

AN EXAMINATION OF PROVING USING A PROBLEM-SOLVING FRAMEWORK

Milos Savic
Michigan State University

A link between proving and problem solving has been well established in the literature (Furinghetti & Morselli, 2009; Weber, 2005). In this paper, I discuss similarities and differences between proving and problem solving by using the Multidimensional Problem-Solving Framework created by Carlson and Bloom (2005) on Livescribe pen data from a previous study of proving (Savic, 2012). I focus on two participants' proving processes: Dr. G, a topologist, and L, a mathematics graduate student. Many similarities were revealed by using the Carlson and Bloom (2005) framework, but also some differences distinguish the proving process from the problem-solving process. In addition, there were noticeable differences between the proving of the mathematician and that of the graduate student. This study may influence a proving-process framework that can encompass both the problem-solving aspects of proving and the differences found.

Key words: Proof, Proving, Proof construction, Problem solving

Proof and proving are central to advanced undergraduate and graduate mathematics courses, yet there is little discussion in these courses of the proving process behind the proofs presented. Since there is an overlap between proving and problem solving (Furinghetti & Morselli, 2009, Weber, 2005), one might look at the problem-solving literature in order to describe some of the aspects of the proving process. I used the Multidimensional Problem-Solving Framework created by Carlson and Bloom (2005), coupled with a data collection technique (Savic, 2012) specifically aimed at collecting the real-time actions that a prover takes, in order to examine the proving processes of a topologist, Dr. G, and a mathematics graduate student, L. I discuss the adequacies and limitations of their framework for describing the observed proving processes. I also discuss the noticeable differences between Dr. G's and L's proving actions. Finally, I conjecture some educational strategies that might be useful for making some implicit actions in the proving process explicit.

Background Literature

Selden, McKee, and Selden (2010) stated that the proving process "play[s] a significant role in both learning and teaching many tertiary mathematical topics, such as abstract algebra or real analysis" (p. 128). In addition, professors teaching upper-division undergraduate mathematics courses often seem to ask students to produce original proofs to assess their understanding. When producing an original proof, some naïve students might not know where to start or how to handle the proving process. This study may help in designing a proving-process framework, which can then be used a tool by students in their own struggles with the proving process.

Both aspiring and current mathematicians seem to need flexibility in their proving styles in order to be successful in mathematics (Weber, 2004; Iannone, 2009). In the mathematics education literature, there are several analytical tools concerning proof production or the proving process, including "proof schemes" (students' ways of "ascertain[ing] for themselves or persuad[ing] others of the truth of a mathematical observation") (Harel & Sowder, 1998, p. 243), affect and behavioral schemas (i.e., habits of mind that further proof production) (Furinghetti & Morselli, 2009; Douek, 1999; Selden, McKee, & Selden, 2010) and semantic or syntactic proof production (Weber & Alcock, 2004).

One analytical tool for the proving process focuses more on the problem-solving aspect while other aspects, according to the authors, may be "autonomous" in proving. Selden and

Selden (2009) described two aspects of a written proof, the formal-rhetorical part and the problem-centered part. According to the authors:

> The *formal-rhetorical* part of a proof (what we have also referred as the *proof framework*) is the part of a proof that depends only on unpacking and using the logical structure of the statement of the theorem, associated definitions, and earlier results . . . . The remaining part of a proof [is] the *problem-centered* part . . . that *does* depend on genuine problem solving, intuition, and a deeper understanding of the concepts involved. (Selden & Selden, 2013, p. 6)

This problem-centered part can be considered as the part of the proof that uses problem-solving, in the sense of Schoenfeld (1985), who stated that a *problem* is a mathematical task for an individual if that person does not already know a method of solution for that task. Past research has indicated connections between proving and problem solving, usually citing proving as a subset of problem solving. In Furinghetti and Morselli (2009), the authors stated that “proof is considered as a special case of problem solving” (p. 71). Weber (2005) considered “proof from an alternative perspective, viewing proof construction as a problem-solving task” (p. 351).

Polya (1957) described many ways to go about problem solving that have been summarized into four overarching steps: “(i) understanding the problem, (ii) developing a plan, (iii) carrying out the plan, and (iv) looking back.” Schoenfeld (1992), somewhat influenced by Polya, described six processes when doing a problem-solving activity: “read, analyze, explore, plan, implement, and verify” (p. 61). Carlson and Bloom (2005) utilized both Polya’s and Schoenfeld’s ideas in creating their multidimensional problem-solving framework, which I describe in the next section.

Carlson and Bloom’s Multidimensional Problem-Solving Framework

The Multidimensional Problem-Solving Framework described by Carlson and Bloom (2005) has four phases, each with the same four associated problem-solving attributes. The four phases are orienting, planning, executing, and checking. The four associated problem solving attributes are resources, heuristics, affect, and monitoring. The table below (Figure 1), taken from Carlson and Bloom (2005, p. 67), shows the multidimensional aspects of their framework.

Executing • Computing • Constructing	Conceptual knowledge, facts and algorithms were accessed when executing, computing and constructing. Without conceptual knowledge, monitoring of constructions was misguided.	Fluency with a wide repertoire of heuristics, algorithms, and computational approaches were needed for the efficient execution of a solution.	Intimacy with the problem, integrity in constructions, frustration, joy, defense mechanisms and concern for aesthetic solutions emerged in the context of constructing and computing.	Conceptual unde numerical intuiti employed to refl sensibility of the progress and pro constructing solu statements.
Checking • Verifying • Decision making	Resources, including well-connected conceptual knowledge informed the solver as to the reasonableness or correctness of the solution attained.	Computational and algorithmic shortcuts were used to verify the correctness of the answers and to ascertain the reasonableness of the computations.	As with the other phases, many affective behaviors were displayed. It is at this phase that frustration sometimes overwhelmed the solver.	Reflections on th correctness and a quality of the sol useful feedback t

Figure 1: Carlson and Bloom’s Multidimensional Problem-Solving Framework (2005, p. 67)

Below I describe each phase, as well as the problem solving attributes associated with each phase. All phases and attributes by Carlson and Bloom (2005) emerged during their analysis of the problem-solving processes of eight research mathematicians and four Ph.D. candidates. An example of one of the problems posed in their study was

Problem 1: A square piece of paper $ABCD$ is white on the front side and black on the back side and has an area of 3 in.[2] *Corner A is folded over to point A' which lies on the diagonal AC such that the total visible area is ½ white and ½ black. How far is A' from the fold line? (Carlson & Bloom, 2005, p. 71)*

Orienting

According to Carlson and Bloom (2005), the orienting phase includes "the predominant behaviors of sense-making, organizing and constructing" (p. 62). Examples of this phase in their study included defining unknowns, sketching a graph, or constructing a table. They stated that an individual may execute these orienting actions with "intense cognitive engagement," ultimately understanding the nature of the problem. Use of resources in the orienting phase can include accessing mathematical concepts, facts, and algorithms. Use of heuristics in the orienting phase can include drawing pictures, labeling unknowns, and classifying the problem. Affect experienced during the orienting phase can include motivation to make sense of the problem, high confidence, and strong mathematical integrity. Finally, use of monitoring in the orienting phase can include self-talk and other reflective behaviors during sense-making, such as asking "What does this mean?"

Planning

Carlson and Bloom (2005) coded a planning phase in a transcript when a participant "appeared to contemplate various solution approaches by imaging the playing-out of each approach, while considering the use of various strategies and tools" (pp. 62-63). In addition, they often observed a subcycle of (a) *conjecture* of a solution, (b) *imagining* what would happen using the conjectured solution, and (c) *evaluating* the validity of that solution during planning phases (See Figure 2).

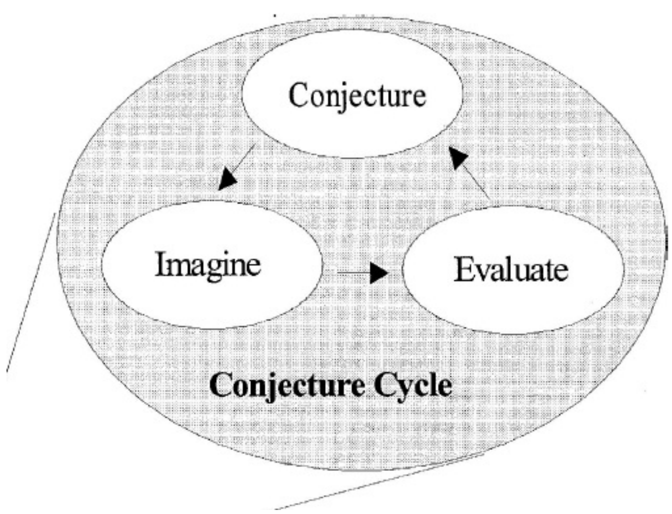


Figure 2: The conjecture-imagine-evaluate [sub]cycle
(Carlson & Bloom, 2005, p. 54)

In Carlson and Bloom's (2005) analysis, this subcycle could be exhibited by their participants either verbally or silently, but the entire planning phase occurred before the executing phase commenced. Resources used during the planning phase included conceptual knowledge and other facts needed to construct conjectures. Heuristics used, if visible to the researchers, included computations and geometric relationships. Affect exhibited by participants during the planning phase included beliefs about the methods or conjectures being employed and about their own abilities to solve the current problem. Monitoring exhibited by Carlson and Bloom's (2005) participants during the planning phase included self-reflection about the effectiveness of their current strategies.

Executing

Carlson and Bloom (2005) noted that the executing phase involved "mathematicians predominantly engaged in behaviors that involved making constructions and carrying out computations" (p. 63). Specific examples included "writing logically connected mathematical statements," using concepts and facts, and using procedures or other computations. Resources used were the same concepts, facts, and procedures that had been used during the prior planning phase. Heuristics used during the execution of the solution included fluency with the algorithms and approaches employed. Affect exhibited in the executing process involved some emotional responses to the attempted solution, such as "intimacy with the problem, frustration, joy, defense mechanisms, and aesthetics in the solution" (p. 67). Monitoring involved the participants having some sensitivity to the progress of their solutions.

Checking

The checking phase was observed when the participants verified their solutions. These behaviors included "spoken reflections by the participants about the reasonableness of the solution and written computations. . .contemplating whether to accept the result and move to the next phase of the solution, or reject the result and cycle back" (Carlson & Bloom, 2005, p. 63). Resources used during the checking phase involved "well-connected conceptual knowledge" for the "reasonableness" of their solutions. Heuristics used included knowledge of "conceptual and algorithmic shortcuts." Affect during the checking phase was similar to other affective behaviors, but frustration might overtake a participant if the solution was incorrect. Monitoring during this phase involved thinking about the "efficiency, correctness, and aesthetic quality of the solution" (Carlson & Bloom, 2005, p. 63).

The cycle of problem solving

Carlson and Bloom (2005) stated that "it is important to note that the mathematicians rarely solved a problem by working through it in linear fashion. These experienced problem solvers typically cycled through the plan-execute-check cycle multiple times when attempting one problem" (p. 63). Carlson and Bloom (2005) also stated that the cycle had an explicit execution, usually in writing, and formal checking that used computations and calculations that were also in writing. All cues exhibited by the participants and observed by the researchers, whether written, verbal, or non-verbal, were used to distinguish between phases.

Research Questions

This connection between problem solving and proof influenced my research questions: Can Carlson and Bloom's (2005) Multidimensional Problem-Solving Framework be used to describe the proving process? If changes or additions are called for, what might they be?

Research Setting

One topologist, Dr. G, and one graduate student, L, were given a set of notes on semigroups and a Livescribe pen and paper, capable of capturing both audio and real-time writing using a small camera the near end of the ballpoint pen. These were two participants of a larger study of nine mathematicians and five graduate students (Savic, 2012), who were asked to answer two questions, provide seven examples, and prove thirteen theorems in the notes. From the first use of the Livescribe equipment for proving or answering all tasks in the notes until the last minute of equipment use, Dr. G totaled five hours and 31 minutes, while L totaled three days, 22 hours, and 11 minutes. I focused the coding of the proving processes on the theorem, "*Theorem 20: A commutative semigroup with no proper ideals is a group.*" From the first use of the Livescribe pen for their proof attempt of Theorem 20 until the last, Dr. G spent three hours and 17 minutes, while L spent 41 minutes. I selected Dr. G's data because he spoke a significant amount of the time while proving and also encountered impasses when proving this theorem. I chose L's data because he was one of only two graduate students who attempted a proof of this theorem and I hoped that his transcript would be amenable to analysis using the Carlson and Bloom (2005) framework. The audio/video recordings were transcribed so that the audio and actions on the paper corresponded. Once the sessions were transcribed, coding was done with the Carlson and Bloom (2005) framework. A sample of the transcription, with coding, can be seen in Table 1.

Table 1: Coding the proving process of Dr. G

Time	Writing/Speaking	Coding
7:02 AM	Th 20: A comm semigp w/ no proper ideals is a gp.	Orienting (Resources)
7:03 AM	Hmm . . . I'm taking a break, breakfast, etc. Back to this later. Must think on this.	Orienting (Resources)
BREAK 7:04 AM - 8:07 AM (Planning)		
8:07 AM	Ok, I thought about this while on a cold walk in the fog. Pf: Given $g \in S$, a semigp., consider the ideal g. . . . (Then he stops and puts "comm." between "a" and "semigp.") . . . $gS = \{gs \mid s \in S\}$. Since S has no proper ideals, $gS = S$, so $\exists g^{-1} \in S \ni gg^{-1} =$ (32 second pause, then he strikes through the whole proof)	Planning (Heuristics) Executing (Resources) Checking (Monitoring) Executing (Resources) Checking (Monitoring)
8:09 AM	First need an identity, not given. (Then he goes back to the expression "$gg^{-1} =$" and writes a question mark with a circle around it.) Turn page.	Checking (Resources), Planning (Resources) Cycling back (Back to planning, re-approaching to answer why there is no identity)

A Description of the Coding of the Sample

At 7:02 AM Dr. G started the proving period of Theorem 20; he had proved the theorems in rest of the notes in the two hours prior to this first attempt. This was his first proving period on Theorem 20. He wrote the theorem on paper, probably orienting himself to what he needed to prove. There was a one-minute pause, which I infer that he was planning or orienting himself to the theorem. Since he had proved Theorems 13-19 quite quickly, his decision to take a walk at 7:04 AM might have been because he had not quickly seen how to attempt this proof. I assume that during the break he might have been planning how to prove Theorem 20, probably using the conjecture-imagine-evaluate subcycle mentioned above. At 8:07 AM, he started executing the idea that he had generated during the walk. He corrected his work by inserting “comm.” to be precise, something that I coded as “checking” and “monitoring” for correctness. Then Dr. G went back to executing his idea, using an element, g, in the semigroup and multiplying it by the whole semigroup to create an ideal. There was a 32-second pause, and then he crossed out the entire proof that he had just written. This was coded as checking. In fact, at 8:09 AM, he wrote why he crossed out this proof attempt: he needed an identity, which had not been given. I coded this as “checking (resources),” because Dr. G apparently used what he knew about groups to verify this attempt. He then cycled back to planning, because there was a 95-minute gap before he wrote something else, beginning with a different idea, and eventually re-orienting himself.

Rules for coding certain situations

For reliability, I asked two other colleagues to code three small excerpts of the transcripts using the Carlson and Bloom (2005) framework. There were certain segments that merited discussion, and we came to an agreement in all of those instances. Using their assistance, I then established a set of rules to help refine the coding process. They were as follows:

1. Both participants had instances in their proving sessions that were pauses in their work. I defined a pause as a period of time in the live data proving session during which the prover does not speak or write. I had asked the participants to prove the theorems at their own leisure with unlimited time, so I was not present to ask them contemporaneously about pauses in their proving. An important aspect of coding a pause would be: If a participant made corrections immediately after a pause, then I would code the pause as “checking.” If after a pause, the participant had an idea or could continue his progress, then I would code the pause as “planning” prior to the executing phase. I also coded participants’ pauses based on what I thought a participant was accomplishing, using my own inferences about their proving process.
    a. When a participant turned off the LiveScribe pen and turned it back on, I considered that a break. All breaks were considered “planning.” This is because almost immediately after a participant turned on the pen after a break, he or she had an idea to try, which is considered in the Carlson and Bloom (2005) framework as “executing.” For example, in Dr. G’s transcript (Table 1), the break from 7:04 AM – 8:07 AM was coded as “planning.”
    b. Many pauses during the proving process were considered “planning,” because of the “executing” phase that occurred immediately afterwards.
2. Speaking was never considered “executing.” Any phase coded as “executing” occurred within the written work, and was only coded this way when it furthered (either correctly or incorrectly) the attempted proof.
3. Any “crossing out” or elimination of any part of the “executing” phase was considered “checking.” An example of this, which occurred at 8:08 AM in Dr. G’s transcript, can be found in Figure 3 below:

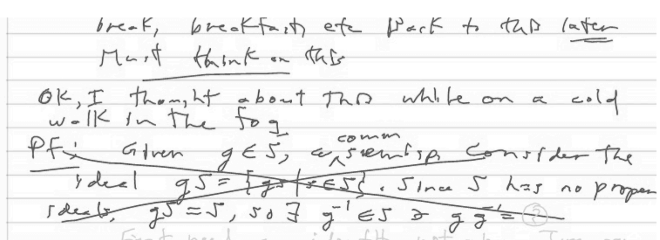


Figure 3: Dr. G’s crossed-out work on Theorem 20

Results

Carlson and Bloom's (2005) Multidimensional Problem-Solving Framework aligned well with most of what the two participants (Dr. G and L) did during the proving process. Using the phases (Orienting, Planning, Executing, and Checking) and the problem solving attributes (Resources, Heuristics, Affect, and Monitoring), I coded all of both transcripts that pertained to Theorem 20 and analyzed the situations that agreed and that disagreed with those in Carlson and Bloom's (2005) framework.

Instances of agreement with Carlson and Bloom's framework

For most portions of the transcripts, the Multidimensional Problem-Solving Framework could be used to code and describe the proving process. There were multiple situations in both transcripts that involved both the planning subcycle (conjecturing, imagining, evaluating) and the larger cycle of planning, executing, and verifying.

The planning subcycle

In Dr. G's spoken discussion of Theorem 21, "If K is a minimal ideal of a commutative semigroup S, then K is a group," he demonstrated the planning subcycle (described in the "Planning" section along with Figure 2) seen in Table 3:

Table 3: An example of the Conjecturing-Imagining-Evaluating subcycle

Time	Writing	Speaking	Coding
9:51 AM		If it [a semigroup S] has a zero element, then that [the zero element] will be a minimal ideal.	Orienting (Resources)
[Cycle Starts		Does that make it a group?	Planning (Conjecturing)
Here]		(Silence for 13 seconds)	Planning (Imagining)
		Well no,	Planning (Evaluating)
		what about the non-negative integers?	Planning (Conjecturing)

The subcycle starts with the question, "Does that make it a group?" which is a conjecture by Dr. G based on the statement, "If it [S] has a zero element, then that will be a minimal ideal." He then paused for 13 seconds. This was coded as "imagining", because Dr. G was imagining what would happen with his conjecture that the existence of a zero element, 0, forces a minimal ideal, namely $\{0\}$, to exist. The next words said by Dr. G were, "Well, no…" acknowledging that he had evaluated where he had been expecting to go with his $\{0\}$ ideal counterexample. I coded that as part of the "evaluating" phase. Finally, he ended this subcycle by conjecturing something about the non-negative integers, completing the planning subcycle. In fact, in the exit interview conducted after his proving sessions, Dr. G acknowledged that he misread the statement and was trying to prove that S was a group. Nonetheless, this episode exhibited the planning subcycle.

Example of a full cycle of planning-executing-checking

In L's proof of Theorem 20, he demonstrated the full planning-executing-checking.**}** cycle seen in Table 4. L did not speak during his proving process, so I conjectured the phases using only his written work.

Table 4: An example of the Planning-Executing-Checking cycle

10:19 AM	First we want to show S has an identity 1. (pauses for 45 sec)	Planning (Heuristics) Planning (Cycling)
10:20 AM	(pauses for 20 sec) If possible. Suppose S has no identity. Then for every $a \in S, ab \neq a$ for all	 Planning (Monitoring) Executing (Heuristics) Executing (Resources)
10:21 AM	(pauses for 25 sec, then crosses out "Then for every $a \in S, ab \neq a$ for all") Let $a \in S$. Let $A = \{ab: b \in S, ab \neq a\}$.	Checking (Monitoring), Planning Executing (Resources)

In this excerpt from L's transcript, he started the proof of Theorem 20 at 10:19 AM with "We want to show S has an identity 1." Since L was writing the sentence to tell himself of his intentions regarding the proof, this statement was coded as "planning". He then paused for a minute and five seconds, and this seemed as if he were going through the conjecturing-imagining-evaluating subcycle reflecting on how he might prove the theorem. The next statement after this pause was, "If possible." Thus, I coded this as a "planning" phase because L was changing how he wanted to approach the proof, and this required him to make sense of which proof framework he would use. After this, he wrote, "Suppose S has no identity." Since L was attempting to prove the theorem, instead of writing guiding sentences as he had done previously (at 10:19 AM), this was coded as "executing." His next sentence, "Then for every $a \in S, ab \neq a$ for all," was coded as "executing" as well. L then paused for 25 seconds. I conjectured that L was considering what he had written, so this pause was coded as "checking". At this point, he crossed out his previous work, and had another idea (namely, creating an ideal) to work with, so I coded this as "planning" right before he executed his idea, "Let $a \in S$. Let $A = \{ab: b \in S, ab \neq a\}$." Hence, the planning-executing-checking cycle can be deemed to have occurred.

Instances of difference with Carlson and Bloom's framework

Cycling back to orienting

Dr. G, after an incubation period (8:09 AM - 9:44 AM) was at a quandary about how to proceed with the proof of the theorem, and in fact had to *reorient* himself to the truth of the theorem. This can be seen in Table 5.

Table 5: An example of reorienting

Time	Writing	Speaking	Coding
8:09 AM	First need an identity, not given. (Then he goes back to the expression "$gg^{-1} =$" and writes a question mark with a circle around it.) Turn page.	None	Checking (Resources) Checking, Planning
BREAK 8:10 AM - 9:44 AM (Planning)			
9:44 AM	Later. I'm suspicious that this is true. Why should the nonexistence of proper ideals force existence of an identity?		Planning (Affect, Monitoring)

	But I don't know many examples, so I don't see a counterexample. (Silence for a minute, followed by ruffled papers, then silence)		Orienting (Resources)

After approaching the proof using a direct proof technique, Dr. G apparently thought about his approach during the incubation period. His next statement was "I'm suspicious this is not true." After this declaration, he claimed that he didn't know many examples. Generating examples for the statement of a theorem can be a way of orienting oneself to the problem of deciding on an approach for proving a theorem. But Dr. G had oriented himself once before (7:02 AM - 8:07 AM) and had already gone part way into the planning-executing-checking cycle. This is an example of an instance in proving when a prover must *reorient* himself in order to reconsider all of the information given. In fact, during his exit interview, Dr. G stated that, "as with many things, at first I thought, 'how can I prove this?' and I didn't immediately think of a proof, so I think, 'what about a counterexample?'" Dr. G was not afraid of reorienting himself with the theorem in order to create counter-examples.

Not completing a full cycle of planning-executing-checking

L finished the proof of Theorem 20 without going through the final checking phase of problem solving. This is displayed in Table 6.

Table 6: An example of not doing the final checking

Time	Writing	Coding
10:48 AM	(pauses for 25 sec, the writes next to $b^{-1} \in S$ from 10:45 AM, "For contradiction", then turns back the page, then pause for 25 sec)	Planning (Resources)
10:49 AM - 10:51 AM	Let $B = \{b \in S: b \text{ has no inverse}\}$. Then $B \neq \emptyset$ because $b' \in B$. So B is a proper ideal of S which is a contradiction. So every element of S has an inverse. (B is proper because $1 \notin B$). Hence S is a group.	Executing (Resources)

After he had written at 10:51 AM "Hence S is a group," he immediately proceeded to the next theorem (Theorem 21). In proving Theorem 20 (incorrectly) L had gone through the planning and executing phases, but had not performed the checking phase. There are two conjectures that I have about this. One conjecture is that he had used a technique that was very close to a technique that he had used previously when trying to prove that S has an identity, and he had taken considerable time and writing (10:39 AM – 10:45 AM) to check his work on that. The other conjecture is that he knew the end of the notes was approaching and he wanted to finish them quickly. This would be understandable, especially since he was a research Ph.D. student that had already successfully completed his comprehensive examinations, so he would probably have preferred working on his own research over proving unrelated theorems.

Discussion

Successes and limitations with the coding

The four phases of the Carlson and Bloom (2005) framework were relevant to the proving process. At first glance, the two participants (Dr. G and L) were always in one of the phases

(Orienting, Planning, Executing, and Checking) during their entire proving sessions for Theorem 20. This suggests that the four phases are very important for the proving process. This further suggests that an expansion of Carlson and Bloom's (2005) framework could potentially provide the mathematics education community a proving-process framework, complete with additional problem-solving attributes that a prover experiences. Some additional problem-solving phases may include incubation, re-orientation, and instances of multiple phases (Checking and Planning) occurring during a pause in the midst of proving or a break from proving in the proving process.

There may also be refinements of the Checking and Planning phases for a future proving-process framework. There were instances of *local planning* or proceeding on only a small part of a proof, and global planning or approaching a proof with a certain framework (Selden & Selden, 1995). An example of local planning was when Dr. G went through the planning subcycle in Table 3, where he asked himself whether a semigroup he had created was a group. An example of *global planning* was when L considered whether to prove Theorem 20 directly or by contradiction in Table 4 (10:19 AM – 10:20 AM). Checking could have also been split into *local checking* (e.g., finding minor errors) and *global checking* (i.e., seeing if a proof attempt is sound). For example, local checking occurred in the middle of Dr. G's first proof attempt when he stopped executing to write "comm." between "a" and "semigroup" (Table 1). Global checking happened about a minute later, where he crossed out his entire proof attempt (seen in both Table 1 and Figure 3).

Carlson and Bloom's (2005) framework describes the process of problem solving well. They provide ample examples from their study that support their framework. When posed a problem like those that Carlson and Bloom (2005) posed in their study (Problem 1 in the Multidimensional Problem-Solving Framework section), mathematicians can rather easily and quickly get conversant with the constraints (orienting) and then go about solving the problem (planning-executing-checking).

However, in my study, the mathematicians were given a theorem (Theorem 20), and had to go about orienting themselves. Some mathematicians (e.g., Dr. G) executed their ideas (e.g., about modifying the hypotheses) early to see where they might lead, but then had to look at the theorem again to analyze why the hypotheses needed modifying. In fact, if one assumes that a statement (in this case, a theorem) could be true or false, one must orient oneself either for a proof or for a counterexample. In mathematical problem solving, unless a problem is posed as a true or false question, some problems can often implicitly be assumed to have a solution.

Carlson and Bloom (2005) audiotaped the mathematicians in their study while they were solving the problems, and were in the room to take notes and answer questions. In my study, however, participants were given a set of notes with unlimited time and not much direction. I was unable to observe their non-verbal actions, something that I conjecture provided Carlson and Bloom (2005) considerable help with their coding. This was a limitation of my study. On the contrary, I was able to capture incubation periods and insight, which were not accounted for in the Carlson and Bloom (2005) framework. This influenced my coding. Breaks are a crucial part of creativity and problem-solving for mathematicians (Savic, 2012), yet are not considered in Carlson and Bloom's (2005) problem-solving framework.

Observed differences between the mathematician and the graduate student

When analyzing all participants in the data collection (nine mathematicians and five graduate students), I found that coding the proof attempts on Theorem 20 would give the best comparison of how one attempts a proof. Six of the nine mathematicians experienced impasses when attempting a proof of Theorem 20, but only two out of five graduate students even attempted a proof. The proof of Theorem 20 was not trivial, which provided a nice comparison between the attempts of Dr. G and L. Notice that Dr. G analyzed situations dealing with the theorem, such as “Why should the nonexistence of proper ideals force existence of an identity?” Dr. G often questioned the constraints of the hypotheses of the theorem. He went a step further and even thought that he might be able to construct a counterexample. According to my coding, L oriented himself at the beginning of the proving period for Theorem 20, and did not question the truth of the theorem, nor the constraints given. My conjecture is that the mathematician (Dr. G) has had substantial experience both with conjecturing his own theorems and adjusting the precise wording of those theorems after attempting unsuccessfully to prove them. He must have had to reorient himself rather often when engaging in mathematical research.

Another observable difference between Dr. G and L was how each handled the “checking” phase. With L, most checking phases were incorporated with the planning phase, where after multiple pauses during the proving of Theorem 20, he both crossed out a certain amount of his previous proof attempt and immediately proceeded to move to the “executing” phase. There was not an observable mixture of phases in Dr. G’s proving process. This may have been due to the amount of speaking Dr. G did, which helped separate the planning and checking phases. But the more important aspect of Dr. G’s checking phase was that he tried to make sense of his failed proving attempts. For example, in Table 1, Dr. G stated, “First need an identity, not given.” He noticed that his previous proving attempt required an identity element, so he must adjust his next proving attempt to accommodate that requirement. During L’s proving attempts, he would make minor adjustments after each attempt but had the same main idea (supposing there is no identity) in mind. Global checking may have helped L gain more information from his proving attempts.

Future Research

It would be an accomplishment if there could be a proving-process framework, similar to Carlson and Bloom’s (2005) problem-solving framework. Such a framework would be helpful in assessing a student’s proving and their phases or problem-solving attributes that need improvement. A teacher could isolate the phases (Orienting, Planning, Executing, and Checking) or problem-solving attributes (Resources, Heuristics, Affect, and Monitoring) that need to be worked on, and focus instruction on that phase/attribute. Also, such a framework might allow researchers to analyze their proof data to analyze and describe new phenomena that they observe.

Additionally, a data collection technique that could capture more phases and attributes would help in developing a proving-process framework. My study gathered written data in real time with synchronized audio. There was no collection of gestures, including when the participants viewed the notes to orient themselves to the statement of a theorem or to gather ideas during a planning phase. In Carlson and Bloom’s (2005) study, the participants were in an interview room for a more-or-less fixed time working continuously on the problems posed. Because their participants had no time for a break or other distracting activity, their data collection technique might have influenced their participants’ creativity. A combination of the two data collection techniques (LiveScribe pen and videoed interview sessions), would be much more informative, but would take more time and resources.

Finally, LiveScribe pens along with Carlson and Bloom’s (2005) framework might assist students in considering the implicit actions of the proving process. For example, students could do either homework or a test with a LiveScribe pen and turn in the pen with their assignment. The professor or a graduate assistant could then upload the data and make a “movie” of a particular part of the student’s proving process. Then the assignment for the students would be to code the movie using the framework. The purpose would be to expose students to certain phases and attributes, thus hoping to make the students mindful of those phases/attributes in the future. In particular, one could see that some students do not incorporate the checking phase in their problem solving or proving. An example of this phenomenon occurred with L's work in Table 6, when he finished his proof of Theorem 20 without any checking. Making phases explicit might help undergraduate students make the transition to proof-based classes quicker, thus shifting the focus of those initial proof-based courses towards their principal purpose of illuminating content.

References

Carlson, M., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent problem-solving framework. *Educational Studies in Mathematics, 58*, 45-75.

Douek, N. (1999). Some remarks about argumentation and mathematical proof and their educational implications. In I. Schwank (Ed.), *Proceedings of the First Conference of the European Society for Research in Mathematics Education, Vol. 1* (pp. 125-139). Osnabrueck: Forschungsinstitute fuer Mathematikdidaktik.

Furinghetti, F., & Morselli, F. (2009). Every unsuccessful problem solver is unsuccessful in his or her own way: Affective and cognitive factors in proving. *Educational Studies of Mathematics, 70*, 71-90.

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.

Iannone, P. (2009). Concept usage in proof production: Mathematicians' perspectives. 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* (pp. 220-225). Taipei, Taiwan: National Taiwan Normal University.

Polya, G. (1957). *How to solve it: A new aspect of mathematical method.* Garden City, NJ: Doubleday.

Savic, M. (2012). What do mathematicians do when they reach a proving impasse? In S. Brown, S. Larsen, K. Marrongelle, & M. Oehrtman, *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 531-535). Portland, OR: Online at http://sigmaa.maa.org/rume/crume2012/RUME_Home/RUME_Conference_Papers_files/RUME_XV_Conference_Papers.pdf.

Schoenfeld, A. H. (1985). *Mathematical problem solving.* Orlando, FL: Academic Press.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 334-370). New York, NY: Macmillian.

Selden, A., & Selden, J. (2013). Proof and problem solving at university level. In M. Santos-Trigo, & L. Moreno-Armella (Eds.), *The Mathematics Enthusiast, Special Issue: International Perspectives on Problem Solving, Research in Mathematics Education, Vol. 10, 1&2* (pp. 303-334).

Selden, A., McKee, K., & Selden, J. (2010). Affect, behavioural schemas and the proving process. *International Journal of Mathematical Education in Science and Technology, 41*, 199-215.

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

Selden, J., & Selden, A. (2009). Teaching proof by coordinating aspects of proofs with students' abilities. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 339-354). New York, NY: Rutledge.

Weber, K. (2004). A framework for describing the processes that undergraduates use to construct proofs. In M. J. Hoines, & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education (PME), Vol. 4* (pp. 425-432). Bergen, Norway: Bergen University College.
Weber, K. (2005). Problem-solving, proving, and learning: The relationship between problem-solving processes and learning opportunities in the activity of proof construction. *Journal of Mathematical Behavior, 24*, 351-360.
Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*, 209-234.

UNDERSTANDING ABSTRACT OBJECTS IN THE CONTEXT OF ABSTRACT ALGEBRA CONCEPTS

Anna Titova
Becker College

ABSTRACT: *This study discusses various theoretical perspectives on abstract concept formation. Students' reasoning about abstract objects is described based on a proposition that abstraction is a shift from abstract to concrete. Existing literature suggested a theoretical framework for the study. The framework describes a process of abstraction through its elements: assembling, theoretical generalization into abstract entity, and articulation. The elements of the theoretical framework are identified from students' interpretations of, and manipulations with, elementary abstract algebra concepts, including the concepts of binary operation, identity, and inverse element, group, and subgroup.*
To accomplish this, students participating in the abstract algebra class were observed during one semester. Analysis of interviews and written artifacts revealed different aspects of students' reasoning about abstract objects. Discussion of the analysis allowed formulating characteristics of processes of abstraction and generalization. The study offers theoretical assumptions on a students' reasoning about abstract objects. The assumptions, therefore, provide implications for instructions and future research.

KEYWORDS: Abstraction, Generalization, Abstract Algebra, Group Theory

Introduction

Abstract thought is considered to be the highest accomplishment of the human intellect as well as its most powerful tool (Ohlsson, Lehitinen, 1997). Even though some mathematical problems can be solved by guessing, trial and error, or experimenting (Halmos, 1982), there is still a need for abstract thought. There is support (Ferguson, 1986) for the hypothesis that abstraction anxiety is an important factor of mathematics anxiety, especially concerning topics which are introduced in the middle grades. By understanding an abstract concept formation we will be able to help students to overcome this anxiety.

This paper presents results of the exploration of the process of abstraction and gives a description of its components and outcomes. The goal is to understand the nature and acquisition of abstraction, so we can help students to bridge the gap from the abstract to concrete. Qualitative approach has been used. Analysis of students' concept formation (knowledge of abstract/mathematical object) is consistent with the tradition of a grounded theory (Charmaz, 2003; Glaser & Strauss, 1967). The study was conducted in the content of group theory.

Theoretical Framework

Piaget (1970a, 1970b) considers two types of cognition: association and assimilation, stating that assimilation implies integration of structures. Piaget distinguishes three aspects of the process of assimilation: repetition, recognition and generalization, which can closely follow each other. In his papers about advanced mathematical thinking, Dubinsky (1991a, 1991) proposes that the concept of reflective abstraction, introduced by Piaget, can be a powerful tool in the process of investigating mathematical thinking and advanced thinking in particular.

In the late 1980s Ed Dubinsky and his colleagues (Clark at al., 1997) started to develop a theory that describes what can possibly be going on in the mind of an individual when he or she is attempting to learn a mathematical concept. In recent years, the mathematics education community at large started to work on developing a theoretical framework, and a curriculum for undergraduate mathematics education. Asiala (Asiala et al, 1996) reported the results on their work; based on the theories of cognitive construction developed by Piaget for younger children, Dubinsky and his colleagues proposed the APOS (action – process – object – schema) theory. A number of studies on topics from calculus and abstract algebra (Zazkis & Dubinsky, 1996; Dubinsky at al, 1994; Brown at al, 1997; etc) were reported using this framework.

The theoretical approach, described by Davydov (1972/1990), is highly relevant to educational research and practice. His theory seems incompatible with the classical Aristotelian theory, in which abstraction is considered to be a mental shift from concrete objects to its mental representation – abstract objects. By contrast, for Davydov, as well as for Ohlsson, Lehitinen (1997), Mitchelmore, White (1994, 1999), Harel and Tall (1991, 1995), abstraction is a shift from abstract to concrete. Ohlsson and Lehitinen provide us with historical examples of scientific theories development; Davydov also gives historical examples and, at the same time criticizes the empirical view on instruction by claiming that empirical character of generalization may cause difficulties in students' mathematical understanding.

Following Piaget (1970 a), the framework for this study considers the process of abstraction as a derivation of higher-order structures from the previously acquired lower-order structures. Moreover, the two types of abstraction are distinguished. One of these types is simple or empirical abstraction – from concrete instances to abstract idea. The second type then is more isolated from the concrete. Davydov (1972/1990) calls this type of abstraction "theoretical abstraction". Theoretical abstraction, based on Davydov's theory, is the theoretical analysis of objects (concrete or previously abstracted) and the construction of a system that summarizes the previous knowledge into the new concept (mathematical object), so it is ready to be applied to particular objects. This abstraction appears from abstract toward concrete and its function is the object's recognition. According to present research, the second type of abstraction is commonly accepted as essential in the process of learning deep mathematical ideas. Similarly, there are two types of generalization – generalization in a sense of Ohlsson and Lehitinen perspectives (which coincides with empirical perspective, described by Davydov), and theoretical generalization. Theoretical generalization is the process of identifying deep, structural similarities, which in turn, identify the inner connections with previously learned ideas. The process of theoretical abstraction leads us to the creation of a new mental object, while the process of theoretical generalization extends the meaning of this new object, searching for inner connections and connections with other structures.

In summary, the genesis of new abstract idea looks like following: (0) initial abstractions; (1) grouping previously acquired abstractions (initial abstractions in a very elementary level); (2) generalization to identify inner connections with previously learned ideas; (3) the shift from abstract idea to a particular example to articulate a new concept. Note that at some level of cognitive development initial abstractions become obsolete since sufficient more complex and concrete-independent ideas are already acquired. The result of this genesis is a new structure which is more complex and more abstract compared to the assembled ideas. Hence, we have hierarchical construction of knowledge, where the next idea is more advanced than the previous one. Moreover, cognitive function of abstraction (from now on, abstraction and generalization are theoretical abstraction and generalization, as defined above) is to enable the assembly of previously existed ideas into a more complex structure. The main function of abstraction is recognition of the object as belonging to a certain class; while construction of a certain class is the main function of generalization, which is making connections between objects:

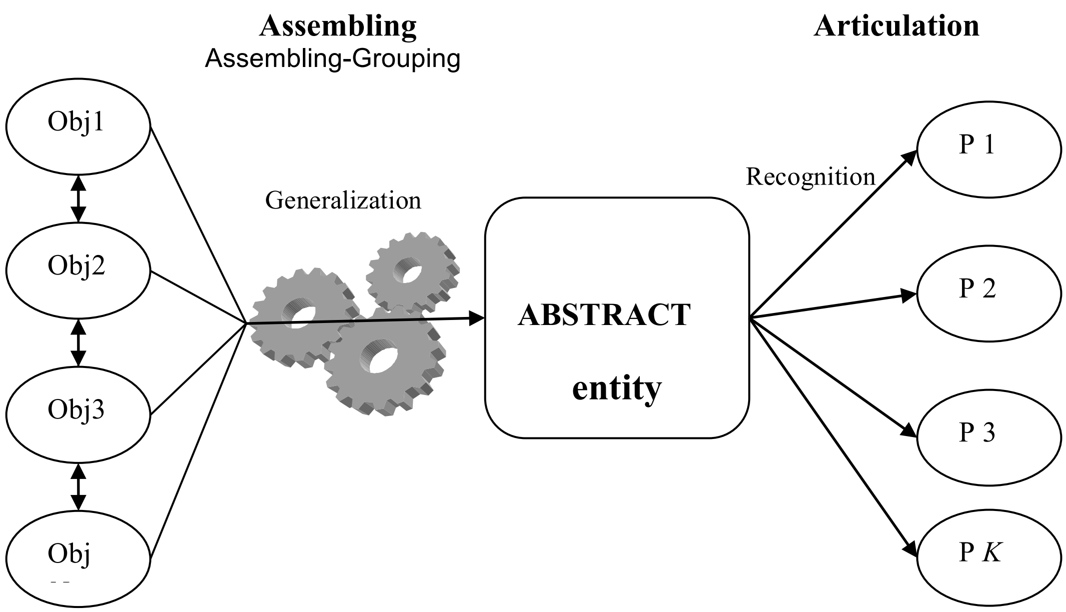


Figure 1: Process of Abstraction

The framework suggests the design of the study and helps to ground the methodology and data collection.

Research Questions

The following questions were formulated based on the theoretical framework:

- What are the main characteristics of the cognitive processes involved in the development of students' understanding of group theory concepts?
- What notions and ideas do students use when they recognize a mathematical object, and why?
- What are the characteristics of students' mathematical knowledge acquisition in the transition from more concrete to more theoretical problem solving activity?

Methodology

To answer the questions above, 22 students, participating in undergraduate Abstract Algebra course were observed during class periods during one semester. The students were mostly math majors and students planning to obtain a secondary math teaching certification. Written assignments (quizzes, homework, exams) were collected from all participants. A group of participants (7 students) was interviewed three times during the semester.

Discussion and Findings

To answer the research questions, students' actions and reasoning about abstract algebra concepts during problem solving activities have been analyzed. The detailed analysis of the data provided answers to the research inquiry.

Results of the study were derived from two general types of analysis: global analysis of students' symbolical and verbal descriptions of mathematical objects and conceptual analysis of the ways in which these concepts were applied to problems. At the initial stage students' interviews were analyzed in chronological order. It allowed forming several preliminary categories that guided further analysis. Students' written work was analyzed afterwards to support the categories and to form new ones. Written work is qualitatively different from verbal responses. Students tried to be more careful and more rigorous when writing down their solutions. It can be explained simply by the fact that the written work was aimed to evaluate students' progress in the course. It was also observed that even if a student did not know the exact definition or a way to solve the problem he or she always tried to give some response. These rather spontaneous responses provided a significant source of information about students' concept formation. During data gathering and analysis the concentration was to look for those responses that showed students' understanding of a concept rather than students' level of preparation. For instance, sometimes students gave a correct definition of a concept but later did not use it in the problem solving process or used it in an incorrect way. After a detailed analysis of students' written artifacts the interviews were analyzed again to support the new assumptions.

Problematic issues that have been found during my analysis can be summarized into several main categories: 1) a correspondence between a set and an operation defined on the set; 2) properties of sets, operations, structures, or elements; 3) use of properties of concrete objects for general conclusions; 4) understanding of abstract algebra statements involving quantifiers; 5) use of definitions – algebraic part versus structural part. The following part is discussing these categories in terms of the theoretical assumptions suggested by the theoretical framework that guided this study. Following the theoretical framework, an abstract object is formed via assembling previously abstracted ideas into a new, more advanced concept. Further, the main function of abstraction is recognition, and during this process the concept is articulated and it leads to the formation of the abstract idea.

The data showed that often the question of a result of a certain operation belonging to the set was not considered. The data also suggested that it was not a matter of forgetting the definition of a binary structure. In most cases, students' responses to tasks such as "Prove (or disprove) that a certain structure is a binary structure" were more or less accurate. Nevertheless, if the question did not specifically ask to check whether a certain element belongs to the structure, then this part was often omitted. In some cases the conclusion about closure was presented in the solution but was based on a different operation.

It looks like the term "binary structure" and the notation $(S, *)$ normally used to represent a binary structure is usually understood by students as a mathematical object with two entrees: a set and an operation. The term and notation do not imply any necessary correspondence or relations between them. Dubinsky et al. (1994) discussed this problem analyzing students' understanding of groups and their subgroups. The study proposed that there are two different visions of a group: 1) a group as a set; and 2) a group as a set with an operation. Similarly for a subgroup: 1) a subgroup as a subset; and 2) a subgroup as a subset with an operation. Analysis of the data collected for this study showed related trends.

Understanding the concept of a binary structure

Binary Operation. Closure. The first connection of a set to its operation appears when the students learn the concept of closure. That is, a set must be closed under the induced operation. The question is why the concept of closure is difficult and how students overcome the problem. First of all, the idea of a binary operation is not completely novel to the students. They have been working with operations from the very beginning of their mathematical experience. It is very well understood by every student that if you take two elements of a set and perform an operation on them the result is another element. Everything the students had to think about was the accurateness of the result. What happens in abstract algebra is that now sets may not necessarily be closed under their operations, so, it must be checked that the result of an operation performed on any two elements of a set must still be in the set. A concept of closure brings a concept of a binary operation to a conceptually different level of abstraction and the previous experience is not a guarantee of success. The number of mistakes I observed in the data (see for example Figure 1 and Figure 2) suggested that the students still try to assimilate the concept of a binary operation to familiar operations.

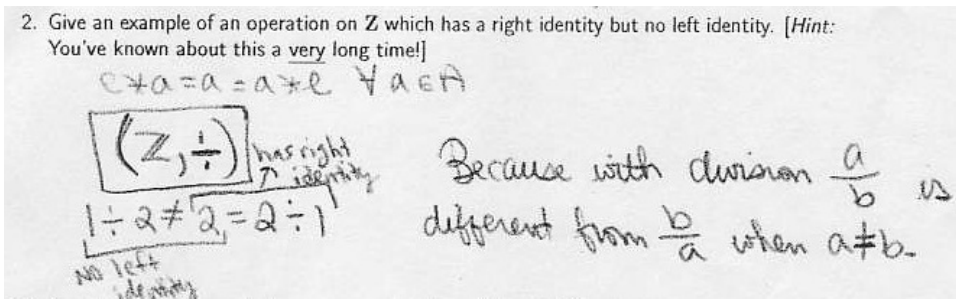


Figure 1. Student's reasoning about division as a binary operation on Z.

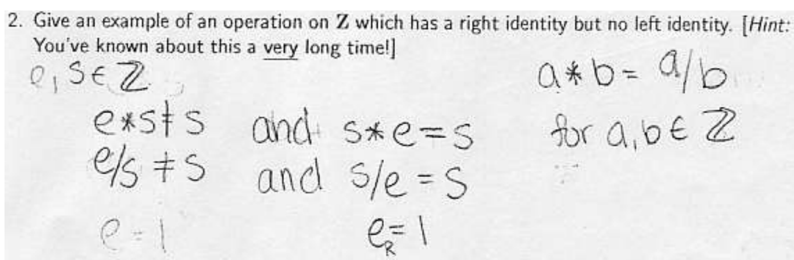


Figure 2. Student's symbolical reasoning about division being binary operation on Z.

Davydov (1972/1990) proposed that the students who experience this problem try to make sense of a binary structure using empirical thoughts (empirical generalization and abstraction). For them a binary operation defined on a set A is a function, where B is some set, not necessarily A. The students assemble ideas of a set, its elements, an operation on any two elements, and a result of the operation on any two elements. By a simple generalization process they develop a simple abstract idea or, in other words, there is a shift from concrete operations (such that addition or multiplication, for instance) to abstract (such as operation "star" defined on set $\{a, b, c\}$). However, often students overlook one important and conceptually different connection among previously known sets and operations and take the definition of an operation on a set for granted. This connection is the link between the result of the operation and the original set. Realization of this connection would provide the concept of an operation with more structural meaning. This process can be referred as theoretical abstraction since there is a derivation of higher-order structures from the previously acquired lower-order structures (Piaget, 1970 a).

Thus, often the process of understanding a binary operation is empirical rather than theoretical. The data provided evidence for the failure of empirical thought about binary operation during the object recognition stage. For instance, when answering the following question: "Give an example of an operation on $\boldsymbol{Z}$ which has a right identity but no left identity", the students often said that division is this type of operation on $\boldsymbol{Z}$ (Figure 1). Indeed, division is not defined on $\boldsymbol{Z}$, since $\boldsymbol{Z}$ is not closed under division and division by 0 is undefined. However, many students recognize division as a binary operation on $\boldsymbol{Z}$ and the reason is that the idea of closure (connection between a result of the operation and a given set) was not a part of students' analysis of the structure ($\boldsymbol{Z}$,). It follows that a theoretical thought is essential in the process of learning deep, structural mathematical ideas.

The interesting phenomenon occurs when students are struggling with assigning a binary operation to a finite subset of the set of integers. Interview 1, question 6, for example:

Question 6. Define a binary operation on $S = \{0,1,2,3,4\}$.

S2: I do not understand what I am supposed to do. Do I need to take two elements from there? OK, so one star two equals three, I guess, if star is addition?

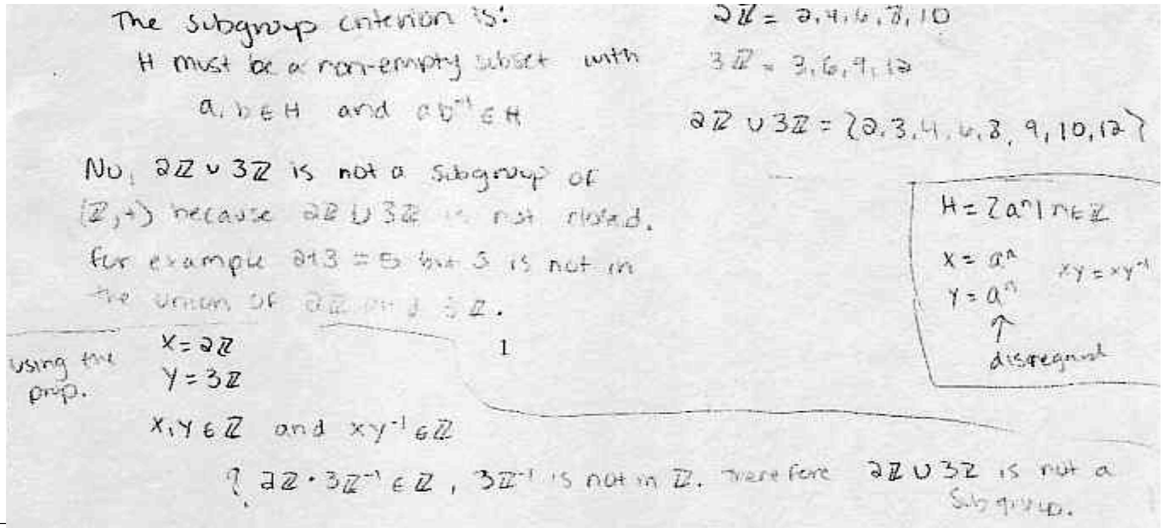
S2: Oh! If I check 2 and 4, I would assign subtraction? Just any operation? Oh I have to assign a binary operation which would work for ALL of them? Ok, well not all of them would work. Like for addition… So I have to think about a binary operation that makes all of those work. Well, I feel like I could have a lot of them! I would have to do something that would say like…a, oh a plus the identity element? Something like that? $a * e = b$ where e is…I mean $a * e = a$, where e is the identity element or something like that.

It does not seem possible for many students to accept the existence of such binary structure. The reason is the closure. Students connect the given finite subset and the operation but for some reason the operation selected is often addition. Most of the students noticed that the subset cannot be closed under addition and concluded that it was impossible to assign a binary operation to the given set or the operation would not be "real":

S2: So, I just have to make up a binary operation, I can't just use addition, multiplication, division or subtraction, cause it's not closed. So I just have to make one up? I feel like there is no one answer. Can it be something like this…Its just taking the first element. It would just be the first element, whatever you are doing. Is it not allowed? It's not a real operation…If you always take the first element and you assign $*$ to, you always get the element from the set.

In this case it appears that students are cognitively placing the objects of recognition to the set of objects for assembling. Operation is the object of recognition, the outcome of the theoretical thinking process. However, the students try to place the concept of operation in the initial stage - assembling process. In responses for question 6 of interview 1, operation of addition is brought to the problem in a rather superficial way, probably due to the elemental association (Halford at al., 1997) of a subset of integers with the binary structure ($\boldsymbol{Z}$, +).

Binary Structures. Group as a set of discrete elements. The understanding of a group as a structure consisting of two objects that interact with each other is complicated and novel for students. The data collected during this study suggests that some students understand a group as a set of elements. The operation in this case does not play an important role in the structure. Some students simply switched from one operation to another.



The subgroup criterion is:
H must be a non-empty subset with
$a, b \in H$ and $ab^{-1} \in H$

No, $2\mathbb{Z} \cup 3\mathbb{Z}$ is not a subgroup of $(\mathbb{Z},+)$ because $2\mathbb{Z} \cup 3\mathbb{Z}$ is not closed.
for example $2+3=5$ but 5 is not in the union of $2\mathbb{Z}$ and $3\mathbb{Z}$.

$2\mathbb{Z} = 2,4,6,8,10$
$3\mathbb{Z} = 3,6,9,12$
$2\mathbb{Z} \cup 3\mathbb{Z} = \{2,3,4,6,8,9,10,12\}$

$H = \{a^n \mid n \in \mathbb{Z}\}$
$x = a^n$ $\quad xy = xy^{-1}$
$y = a^n$
disregard

using the prop.
$X = 2\mathbb{Z}$
$Y = 3\mathbb{Z}$
$X, Y \in \mathbb{Z}$ and $xy^{-1} \in \mathbb{Z}$
? $2\mathbb{Z} \cdot 3\mathbb{Z}^{-1} \in \mathbb{Z}$, $3\mathbb{Z}^{-1}$ is not in $\mathbb{Z}$. Therefore $2\mathbb{Z} \cup 3\mathbb{Z}$ is not a subgroup.

1

Figure 3. Switching addition to multiplication.

It suggests that for the students operation is not an attribute of a binary structure but rather a separate object which can be used or changed if notation suggests. Indeed a^n would always represent multiplicative exponent in Algebra and Calculus.

Another example of students' responses which emphasize understanding of a group as a set occurred when the students were asked to define a General Linear group.

All $n \times n$ matricies with entries from the rational numbers

Figure 4. Student's definition of the General Linear Group *GL (n, Q).*

During the interviews most of the students mentioned both a set and an operation when defining a group. However, only a few students defined a General Linear group *GL(n,* ***Q****)* in terms of a set and an operation. Most of them limit the response to the description of the set of invertible $n \times n$ matrices with entries from ***Q***. Further, in response to the following question from Interview 2: Determine whether {4, 8, 12, 16} is a group under multiplication (mod 20), many students noticed that it cannot be a group since 1 (or 0 in some responses) does not belong to the set. The conclusion is based on the elements of the set, not on the given binary structure:

> S2: That would be yes, it's [generator] 4. If you start with element 4 and mod 20, then yes it is. Well…it needs 0. If you want to say 4 mod 20: 4, 8, 12, 16, and this is…
>
> I: the operation is multiplication

S2: Then it's not. Does not have an identity or inverse. Wait, I am confused with this. It's not a group. Well it's under multiplication it does not have an identity or inverse. Multiplicative identity is 1.

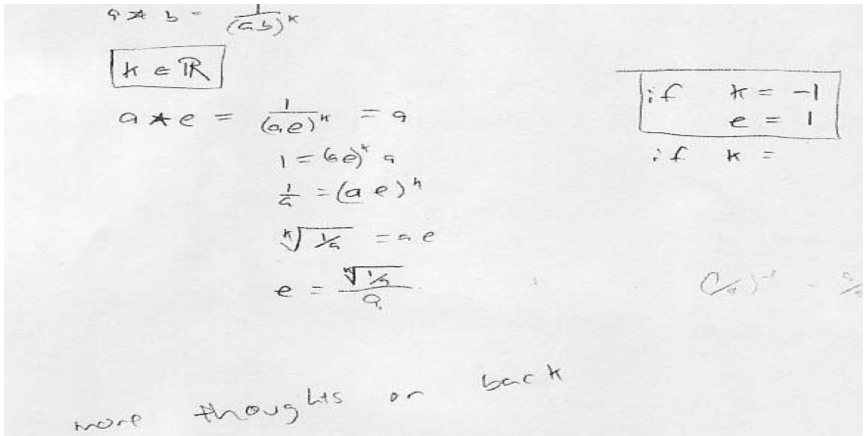
At the early stage of understanding the binary structure concept, students construct their knowledge based on previously learned objects. In order to understand a complex idea such as binary structure, students must have other ideas as parts. So, the elements of a binary structure represent these ideas. The process of generalization initializes connections between the elements, and groups these elements in a set. Thus, the new created abstract entity simply repeats the one that already exists. In this case operation defined on a binary structure is not a part of the assembling process and exists disjointedly from the set. This is the process of generalization in Ohlsson's, Lehitinen's (1997) sense, or Davydov's (1972/1990) empirical generalization. The idea of a group as a set is formed via extraction commonalities from concrete examples, based on visual representations, symbols, discourse, etc. For example, ($\boldsymbol{Z}$, +), ($\boldsymbol{Z}_2$, $+_2$), ($\mathbf{Z}_3$, $+_3$), as group instances, have integer elements in common, 0 as identity element, and based on these examples the idea of a group is empirically generalized. In this case the abstract idea is not complete and further the main function of abstraction fails. According to Davydov (1972/1990), the main function of abstraction is object recognition. So, the abstraction moves from abstract (formed entity) toward concrete (recognition) (Ohlsson S, Lehitinen E., 1997). In this sense recognition is the main function of theoretical abstraction. Some of the students recognized a General Linear group as merely a set of invertible *n* by *n* matrices. Also a structure ($\{4, 8, 12, 16\}\times_{20}$) was not recognized as a group since 1 (or/and 0 in some responses) is not in the set. In both cases the operation is not considered to be a part of the structure and the conclusion is based only on the elements of the set.

In these cases we deal with empirical type of generalization, or simple generalization, using Piagetian (1970) terms. Students assembled examples of groups they studied and extracted commonalities from the sets. According to Piaget (1966, 1970), simple generalization is a part of the empirical abstraction process. In the literature, abstract algebra objects, like many other objects in mathematics, require advanced thinking. Mathematical ideas are complex structures and require a theoretical thought (Davydov, 1972/90). The data illustrated a failure of empirical abstraction to recognize correct objects during problem solving (or working with concrete examples). It follows that the main function of abstraction (recognition) is not supported by empirical abstraction. Theoretical thought requires more than just extracting commonalities from previously learned ideas. Again, in the problem about the set {4, 8, 12, 16} under multiplication mod 20, assuming that students' view on this structure is bounded by the set only, we observe a process of simple generalization as merely a search for commonalities between the given set and other structures which are known to be groups. Element 1 is not in the set, so the structure is not a group. The process of assembling in situations like this is not complete and it causes the ignition of empirical generalization as a replacement for theoretical generalization.

While solving the problem about the set {4, 8, 12, 16}, many students tried to find a generator for the set to prove that this is a cyclic group. This way of thinking is unusual, but correct. However, the students are thinking about the structure without considering a given operation at all. They try to play around with numbers, using the familiar operations of addition or multiplication. Some students at first stated that 4 is a generator since 4 + 4 is 8, 8 + 4 is 12, etc. This demonstrates a misconception caused by assembling wrong ideas into the generalization process. Obviously, the assembled ideas include the ideas of a set, an element, a generator, a cyclic group, addition/multiplication, closure. I think that in this case we deal with the process of theoretical generalization and further with theoretical abstraction. For this group of participants, the structure is recognized to be a group if it is isomorphic (although I do not think that the students really had a thought about isomorphic structures but they obviously had an idea about structures with similar properties: a cyclic group is a group) to a cyclic group (or itself is cyclic). Now the problem is restrained to the following: find a generator for the set, and the process of thinking goes as following: 1) find a generator; 2) if a generator is found, the structure is cyclic; 3) it is a group. It means that the process of theoretical generalization is completed, or the inner connections between the objects of assembling were analyzed. Under this assumption the recognition and conclusion are not at all controversial and we observe all stages of the process of theoretical abstraction. However, the operational part of the thinking process is misleading, since the operation is not standard addition/multiplication. Thus, the initial assembling is not accurate and the conclusion either cannot be drawn, or is drawn incorrectly. Although the operation of addition/multiplication is a part of this assembling process, understanding of groups in this case is the "group as a set" type, since the actual operation, defined on the structure, was not considered, or at least was not considered in connection with the set.

Understanding of a subgroup in general

Since a subgroup is a group itself with some additional conditions, it is difficult to distinguish students' understanding of groups and subgroups. It is clear that some issues in the understanding of a subgroup are caused by the misunderstanding of the concept of a group. Data analysis showed that even if a group or a subgroup is considered to be a binary structure (set together with the operation), the problem of "closure" often persists. Repeatedly students did not consider operation as part of the subgroup concept, or did not connect a group and its subgroup operationally. Dubinsky et al. (1994) suggests that "an individual's development of the concepts of group and subgroup may be synthesized simultaneously" (p. 273). Indeed, in order to learn the concept of a subgroup, the ideas of a group, a subset, and a group operation are assembled and theoretically generalized into an abstract entity for recognition and final concept formation. At the same time, group as a part of the assembling process has a complicated nature. Students must have an idea of a group in assembling not only because a subgroup itself is a group by definition, but also because a subgroup is a substructure of a bigger structure which is also a group. This is a very interesting issue. Assuming that the concept of a group is already learned (previously learned abstract ideas such as set, operation, closure, associativity, identity element, inverse element, etc. are assembled, generalized into an abstract entity and mastered on concrete examples), then the bigger structure which is given to be a group must be recognized as a structure with specific properties which are affecting a substructure. Interestingly, during the interviews some students said that they cannot say anything about G (given to be a group) since they do not know what G is. So, without having a concrete structure the students could not rely on group's axioms and properties. It means that initial assembling was not complete and/or the generalization process was empirical rather than theoretical, and concrete examples played the role of abstract ideas in the assembling process. Concrete examples must be a part of the assembling process. In many cases students needed both abstract ideas and concrete examples to understand a more complex idea:



Back:

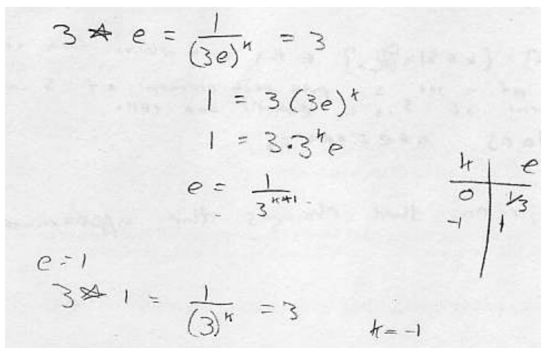


Figure 5. Student's reasoning with concrete examples.

However, the shift from these ideas and objects to the new abstract entity must be a result of theoretical, not empirical, generalization.

The analysis showed that students have difficulty understanding connections between a group and its subgroups, both operational and via element. Student's responses revealed three major misconceptions about the concept of a subgroup. First, for some students understanding of a subgroup is similar to the understanding of groups as sets, but one does not necessarily implies another. For instance, those students who at first understood a group as a set would not necessarily transfer this understanding onto subgroups and vice versa. For some students a group is a set with the operation while a subgroup is just a subset, a part of a bigger structure. A subgroup exists if a subset exists. For example, several students claimed that the set of odd integers is a subgroup of ($\boldsymbol{Z}$, +).

The study has also shown that students have problems seeing structural connections between groups and its subgroups. Sometimes they only comprehend elements connection. Note that this case is different from the one described above. This time students realize that a subgroup is a group itself under an assigned operation. It is not merely a subset of a bigger set, it is a structure. Nevertheless, the assigned operation is not necessarily the group operation. For example, some of the responses defended that ($\boldsymbol{Z}_n$,$+_n$) is a subgroup of ($\boldsymbol{Z}$, +), since it is a group and Z_n is a subset of Z. I also observed a change of the subgroup operation from the group operation to a different operation during problem solving activity (Figure 3).

In addition, I observed responses that not only demonstrate students' understanding of a subgroup as a subset of a given structure but, in addition, a subgroup is understood as a group and a subgroup has the group operation. However, the concept of binary operation is causing difficulty. It is well illustrated by the following response (Figure 6):

Problem 3: Is it possible to find two nontrivial subgroups H and K of (**Z**, +) such that H∩K={0}? If so, give an example. If not, why not?

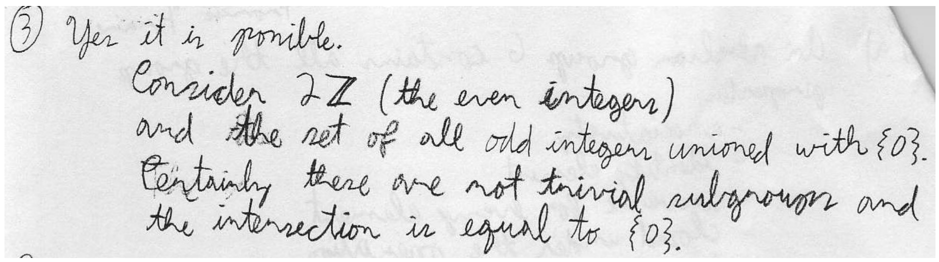

(3) Yes it is possible.
Consider 2$\mathbb{Z}$ (the even integers)
and the set of all odd integers unioned with {0}.
Certainly these are not trivial subgroups and
the intersection is equal to {0}.

Figure 6. Using sets which are not subgroups.

A set of odd integers together with 0 is a subgroup of (Z, +). Element 0 is added to the odd integers set. It suggests that the student understands that the structure has an identity element. It follows from the fact that a subgroup is a group itself. Moreover, it looks like they understand that the operation is addition, since 0 is the additive identity. So, the only problem is the closure of the structure.

In light of theoretical perspectives it seems that in the standard learning sequence "group – subgroup" the concept of a subgroup is the merger of two concepts. One can have an abstract idea of a subgroup only if the idea of a group has already emerged. At the same time a subgroup being a group has special properties which define the subgroup, therefore it is important that students understand these characteristics. It suggests that to construct the abstract idea of a subgroup students need more ideas for assembling than for understanding the concept of groups.

When students are solving problems involving groups and subgroups they are acting upon given objects and operations using abstract ideas they already have. At this stage, the concept of a subgroup is not abstracted yet. It requires more mastery: more concrete examples. At this stage students already have the required minimum of assembled ideas, for instance: set, operation, closure, subset, identity, inverse and group. Still, assembling of various ideas is not enough. Assembling must be followed by theoretical generalization and then articulation. It seems like the concept of a subgroup is not articulated enough. Theoretical generalization as a part of theoretical thought suggests that, for instance, the existence of an inverse element for every element of the set must be proved and is not given for granted.

Definitions of objects. How students use them

The data (both written artifacts and interviews) illustrated that the students mostly used informal definitions of the concepts they study. Normally, an exam or a quiz in the course included questions about definitions, and it was always announced in class before the test so that they could study the definitions. Every quiz and exam was structured in such a way that students had to give a definition of a concept and then solve problems involving this concept. This testing strategy helped to identify some interesting patterns. It showed whether the students used a definition they just formulated, and if they did, in what way and what parts of the definition the students considered being the most important and significant for concept recognition and handling.

The theoretical framework suggests that a definition is the initial stage of concept formation. A definition suggests ideas for assembling. For example: a group is a set, closed under an assigned operation, the operation must be associative, an identity element must be in the set and every element of the set must have an inverse. The definition puts forward some previously abstracted ideas for assembling. Analysis of the connections between the ideas, and articulation follow the assembling. Later, when concepts are being recognized in concrete problems student also must refer to definitions to collect objects from the assembling process, which must be recognized first.

Interestingly, students were annoyed by definition questions during the interviews. One of the interviewees noted that they "just use it" but she/he was not sure how to state the requested definition. Also even if a definition was stated correctly, students infrequently used it in their problem solving; or used some parts of it. It suggests that there is a gap between the abstract entity students have constructed from the definition and the articulation process, the recognition per se. Recalling the assumptions that definitions suggest ideas for assembling, it follows that not all ideas are being generalized into an abstract entity. Further, when recognizing objects using definitions, students simply recognize the objects that were parts of the assembling process. However, some of these objects may not find their place in the abstract entity (and thus will not be recognized) or could be considered as unimportant. It results in failure of the recognition using the definition.

Another observation that follows from responses similar to ones illustrated in Figures 1 and 2 is that students are often concentrated on the algebraic part of the definition, ignoring other conditions that can be stated symbolically or verbally. Equations easily become assembling objects, while the rest of the definition is omitted.

Quantifiers

The study did not intend to explore students' discourse or use of quantifiers. However, as the data showed, this issue plays a big role in students learning process. Some students who participated in the study did not use quantifiers at all when defining objects. Sometimes, missing quantifiers did not mean that the concept was not recognized or used properly during problem solving process. The preliminary analysis of the interviews suggested looking more carefully at the written work in terms of the presence of quantifiers. Students used quantifiers more often when writing statements but sometimes students changed the order of quantifiers they used. For example, instead of writing $\forall\exists$ statement they had $\forall\exists$ statement:

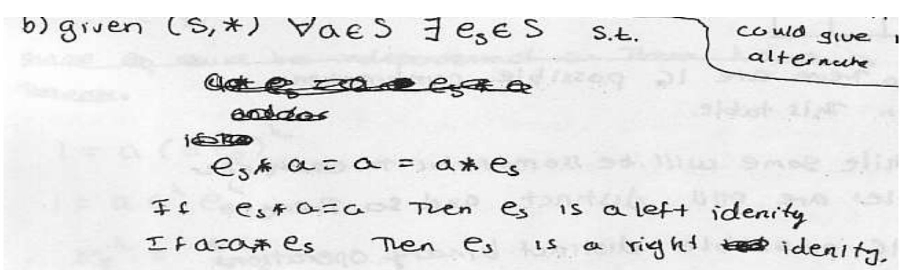

b) given $(S,*)$ $\forall a \in S$ $\exists e_s \in S$ s.t.
could give
alternate
$e_s * a = a = a * e_s$
If $e_s * a = a$ Then e_s is a left idenity
If $a = a * e_s$ Then e_s is a right idenity.

Figure 7. Illustration of $\forall\exists$ – $\exists\forall$ problem in identity definition.

Quantification question is very important for concept formation and requires more exploration.

Conclusions and Implications

The study was aimed to understand how students reason about abstract algebra concepts, how they operate with abstract objects, how abstract concepts are generated in general, and what connections between abstract concepts and concrete examples students see. The study is guided by the theoretical framework that is based on Piaget, Ohlsson, Lehitinen, and Davydov's view on students' reasoning and the processes of abstraction and generalization.

The study showed that one needs to have previously abstracted ideas to understand a new abstract structure. Moreover, data analysis and further discussion ascertained that an abstract concept cannot be learned without concrete examples and problems that involve the concept. In other words the articulation of an abstract concept is required for coherent structure formation. This section summarizes the discussion of the findings and makes several conclusions about students' understanding of abstract objects. The causes of major problems in students' learning that can lead to misconceptions and inability to solve abstract algebra problems are also summarized here.

The data and theoretical framework suggested the model of abstract concept formation – process of abstraction. At the first stage of the learning process students are often given a definition of a concept being studied. Sometimes several simple examples precede the definition. These activities give students a chance to generate a preliminary set of objects for assembling. All these objects are previously learned abstract ideas. The process of assembling is followed up by the process of theoretical generalization. Since a definition usually gives only a preliminary set of ideas for assembling, it is most likely impossible to coherently understand inner connections between the ideas and form a plausible abstract entity. For this reason, this process as described as a preliminary generalization. The next standard instructional step is illustration of the concept via various examples. During this stage students are getting their first articulation experience and make first attempts to concept recognition. The first stage does not necessarily lead to consistent concept formation. A concept said to be generated if it is recognized during problem solving together with all its properties. At this stage a student should be able to exemplify and counter exemplify the concept. It means that when the concept is learned the process of abstraction of these objects gets into the following static form: 1) connected assembled ideas; 2) complete understanding of meaningful inner connections; 3) open-minded recognition of the object. At this stage, students should be able to move easily from object recognition to assembled ideas, if needed. Most of the time, if not always, the first stage does not give the result of the static form described earlier. After the concept was defined and exemplified, instructions usually are followed up by a problem solving activity where the concept that is being studied interacts with other concepts and ideas. This is the second stage of the concept formation process. At this stage students are exposed to additional ideas for assembling, and make more thorough theoretical generalization for the correct recognition of the object. Also this stage provides students with an understanding of properties of the concept and they again could add ideas for assembling. These stages are repeated as many times as needed.

Possible quandaries of a concept formation are coming out of the theoretical configuration described above. It was noticed that all the stages of abstract concept formation are interconnected. There is a constant interaction between processes (assembling and articulation) within the process of abstraction. This observation implies that if there is a problem with one process the abstract concept cannot be appropriately formed. The discussion of the findings led to the following summary of possible predicaments for concept formation: 1) Empirical generalization and abstraction instead of theoretical. Students are trying to learn concepts by extracting commonalities from given concrete objects and examples. 2) Assembling of unsuitable ideas. Students mistakenly assemble some ideas which are not supposed to be assembled to learn a certain concept. As a result, theoretical generalization results in a misleading abstract entity and further in false conclusions which look true under students' arguments. 3) Insufficient number of assembled ideas. 4) Making the object of recognition (during problem solving) one of the ideas for assembling. 5) Insufficient articulation. Students find it difficult to provide examples and especially counterexamples. 6) Isolation of concrete examples from objects of assembling. Sometimes students do not see the interaction between the concrete examples and the abstract structure. A concrete example is considered to be a static object with fixed properties. For instance, students know that $(\boldsymbol{Z}, +)$ is a group with identity 0 and the inverse for any element is its opposite integer, but they do not question it, they simply take it for granted.

Awareness of these predicaments can help to create meaningful instructional activities and classroom settings, giving enough examples and time so that students can articulate the concept they study. The following is summarizing some recommendations for classroom activities and instructions: 1. Writing definitions being used each time during problem solving activities and exemplifying concepts, stressing which part of the definition works for each step of the solution. It helps to compel the appreciation of formal logic and make definitions more explicit. It would allow students to properly use definitions when proving statements. It is especially important for teachers' preparation since NCTM Standards (2000) suggest that students should be able to develop and evaluate mathematical arguments and proofs. 2. Discussions of possible ideas for assembling. Preparation of sets of questions to stress the connections between assembled ideas. 3. Collaborative activities, group discussions where students can put together more ideas for assembling and discuss connections between the ideas. 4. Special attention to "problem posing" activities and special attention to counterexamples. 5. Constant analysis of students' work can suggest unique instructional approach to an individual student as well as a group of students in terms of ideas for assembling and articulation.

Theoretical conclusions of this study can be applied to different mathematical courses at various levels. They are not limited by mathematics only and can be functional for other areas. To elaborate on these predicaments, more exploration, possibly within a different mathematical content, is needed.

REFERENCES

Asiala, M., Brown, A., DeVries, D., Dubinsky, E., Mathews, D. and Thomas, K. (1996), A framework for research and curriculum development in undergraduate mathematics education, *Research in Collegiate Mathematics Education* 2, 1–32.

Brown, A., DeVries, D., Dubinsky, E., Thomas, K. (1997). Learning Binary Operations, Groups, and Subgroups. *Journal of Mathematical Behavior,* 16(3), 187-239.

Charmaz, K. (2003). Qualitative Interviewing and Grounded Theory Analysis. In Holstein, J.A., & Gubrium, J.F. (Eds.) *Inside Interviewing: New Lenses, New Concerns* (pp. 311- 330). Thousands Oaks, CA: Sage.

Clark, J. M., Cordero, F., Cottrill, J., Czarnocha, B., DeVries, D. J., St. John, D., Tolias, T., & Vidakovi´c, D. (1997). Constructing a schema: The case of the chain rule. Journal of Mathematical Behavior, 16, 345–364.

Davydov, V.V, (1972/1990) Types of Generalization in Instruction. Logical and Psychological Problems in the Structuring of School Curricula. *Soviet Studies in Mathematics Education.* Vol. 2 (J. Kilpatrick, J. Teller, Trans) Reston, VA NCTM (Original published 1972, Moscow, Russia)

Dubinsky, E. (1991). Reflective abstraction in Advanced Mathematical Thinking. In Tall, D. (Ed.), *Advanced Mathematical Thinking*. Dordrecht, The Netherlands: Kluwer, 95 – 123.

Dubinsky, E., Dautermann, J., Leron, U. and Zazkis, R. (1994), On learning fundamental concepts of Group Theory, *Educational Studies in Mathematics* 27, 267–305.

Ferguson, R.D. (1986) Abstraction Anxiety: A Factor of Mathematics Anxiety. *Journal for Research in Mathematics Education.* Vol.17, no.2, pp 145-50.

Glaser, Barney G. and Strauss, Anselm L. (1967) *The discovery of grounded theory: strategies for qualitative research*. Chicago.: Aldine

Harel, G.(1995). Abstraction, Scheme Reconstruction, and the Necessity Principle. Unpublished paper. Perdue University.

Harel, G, Tall, D.(1991) The General, the Abstract, and the Generic in Advanced Mathematics. *For the Learning of Mathematics--An International Journal of Mathematics Education*. Vol.11, no.1 pp.38-42

Mitchelmore, M.(1994) Abstraction, Generalization and Conceptual Change in Mathematics. *Hiroshima Journal of Mathematics Education*, v2 p45-57

Mitchelmore, M. C, White, P.(1995) Abstraction in Mathematics: Conflict, Resolution and Application. *Mathematics Education Research Journal*, v7 n1 p50-68

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

Ohlsson S, Lehitinen E. (1997) Abstraction and the acquisition of complex ideas. *International Journal of Educational Research.* Vol. 27, no. 1

Piaget, J.(1970 b). *Genetic Epistemology.* Columbia University Press.

Piaget, J.(1970 a). *The Principals of Genetic epistemology.* London.

White, P, Mitchelmore, M. (1999). Learning Mathematics: a New Look at Generalization and Abstraction. [On-Line]. Available at: http://www.aare.edu.au/99pap/whi99309.htm

Zazkis, R. Dubinsky, E. (1996), Dihedral Groups: A Tale of Two Interpretations, *CBMS Issues in Mathematics Education,* vol. 6.

STUDENTS' MEANINGS FOR RATE IN TWO AND THREE DIMENSIONS

Eric Weber
Oregon State University
Eric.Weber@oregonstate.edu

This paper describes two first-semester calculus students' meanings for rate of change, and how these meanings shaped their conversations and interpretations of rate in three dimensions. I present the theoretical structure of the teaching experiment in which the two students participated, use excerpts from the teaching experiment to illustrate the development of their ways of understanding, and present a retrospective analysis that characterizes the meanings I believed the students possessed. I conclude by discussing the need to use results from literature about generalization and abstraction to inform the study of student thinking in multivariable calculus.

Key words: Rate of change, covariation, functions, graphs, quantitative reasoning.

Background

Models in physics, engineering, economics and mathematics inevitably involve reasoning about change among the many variables in the system. Students must grapple with the meaning of "varying the input" when there are two independent variables and determine rates when there is no "natural" direction in which change happens. These are dilemmas that are not easily explained with the ability to only think of a single-variable function. However, little is known about ways of understanding single-variable functions and their rate of change that are more or less propitious for learning about multivariable functions and how to think about rate of change in the context of multivariable functions. It is therefore important that we understand affordances and constraints in students' thinking in regard to conceptualizing functions of two or more variables and in regard to their conceptualizing rate of change in those contexts.

The purpose of this paper is to propose ways of thinking about rate of change that I hypothesized would support students in extending their understanding of rate from two to three dimensions in a coherent way. I propose a conceptual analysis steeped in quantitative and covariational reasoning that describes productive meanings students should have to think about rate in a way that extends to functions of many variables. I then provide insight into students' ways of thinking and ways of understanding rate of change in the plane and in space as they participated in a teaching experiment that was focused on these ideas.

Assumptions about Meaning

Thompson (in press) traced the development of meaning as a construct throughout the 20th century to suggest that coherent meanings are at the heart of the mathematics that we want teachers to teach and what we want students to learn. This led Thompson to focus on what it means to understand. He relied on Piaget's notion of understanding as synonymous with assimilation to a scheme, where a scheme is an organization of operations and images. Thompson characterized Piaget's meaning for assimilation as similar to imbuing something with meaning, which goes beyond the standard description of assimilation as absorption of information. He concluded that constructing a meaning is the same as constructing an understanding, and that constructing a meaning occurs by repeatedly constructing understandings anew (Thompson, in press). I drew from Thompson's characterization of meaning by focusing specifically on students' assimilations, in which they imbue meaning to something. This does not mean that the meanings students possess are the ones we intend. Instead, it means that I took seriously the effort of representing students' understanding by making inferences about the meanings they possessed and developed.

Foundational Meanings: Quantitative and Covariational Reasoning

The scheme of meanings for rate on which the teaching experiment was based relied on quantitative and covariatonal reasoning. At the heart of understanding rate of change is thinking about rate as a quantification of how fast one quantity is changing with respect to another. This understanding requires that the students construct these quantities, and then imagine them varying simultaneously. The construction of these quantities and their interrelationships relies on quantitative reasoning.

As an example, consider the photo below (Figure 1). Suppose that Persons A, B, and C are looking at the photo, and that Person A notes that there is a crosswalk and a blurred image of a car passing through it. Suppose that Person B notices that there is a "space" between the crosswalk lines, but does not think any further about the "space" being a distance. Person C attends to the distance between the crosswalk lines as a number of feet and notices that the blur in the photo is caused by the car moving some number of feet in the amount of time that the camera's shutter was open. Person A's conception is non-quantitative. Person B's conception might be called *proto*-quantitative, meaning that this way of thinking is necessary for a quantitative conception of the situation but it is not actually quantitative. Person C is thinking quantitatively about the situation presented in the photo by having constructed attributes of the car and the crosswalk and imagined them as measurable. Only Person C is positioned to understand that with two pieces of information (that the shutter was open for 1/20 second and that the crosswalk is 10 feet wide) she can estimate the car's speed: The blur is slightly more than 1/3 the crosswalk's width, so the car traveled approximately 3.4 feet in 1/20 second. This is equivalent to traveling at a constant speed of 68 ft/sec, or about 46 mi/hr.

Figure 1. Conceiving of the speed of a moving car.

Quantitative reasoning is the analysis of a situation in a quantitative structure, which Thompson refers to as a network of quantities and quantitative relationships. If a student is to think about a complicated situation involving three or more quantities and an invariant relationship between those quantities, a dynamic mental image of how those quantities are related is critical. That image positions a student to think about how quantity 1 varies with quantity 2, how quantity 2 varies with quantity 3, and how quantity 1 varies with quantity 3. These images position students to reason covariationally. Saldanha and Thompson (1998) described covariational reasoning as holding in mind a sustained image of two quantities' magnitudes simultaneously as they vary. Saldanha & Thompson's image of covariational reasoning relied on measured properties of objects, and distinguished simultaneous, continuous change from successive, discrete change. They spoke of images of successive, coordinated changes in two quantities as an early form of continuous covariation that, if developed, becomes an image of simultaneous change.

Carlson, Jacobs, Coe, Larsen, and Hsu (2002) described covariational reasoning and accompanying framework built on Saldanha & Thompson's (1998) description of continuous covariation to propose ways of thinking about how those quantities covaried. Carlson suggested that covariational reasoning allows a student to extract increasingly complicated patterns relating x and $f(x)$ from the table of values by ways of thinking the student might use to understand what occurs between those values. Thompson (2011) expanded on his notion of continuous covariation to propose how a student's construction of quantities and their continuous variation could support an image of those quantities' continuous covariation. Thompson introduced the construct of *conceptual time* to propose a plausible scheme of meanings for a student's construction of a quantity that would support an image of that quantity varying with another quantity simultaneously. In essence, Thompson characterized the conceptualization of a quantity's value varying continuously as that its value varies in infinitesimal bits, with the anticipation that within each bit the value varies continuously (Thompson, 2011, p. 45).

Thompson's characterization of variation extends to imagining two quantities covarying, represented here as $(x_e, y_e) = (x(t_e), y(t_e))$, where (x_e, y_e) represents an image of uniting two quantities, and then varying them in tandem over intervals of conceptual time (Thompson, 2011, p. 48). This characterization of covariation is similar to conceiving a function as defined parametrically. If a student has this conception of covariation in mind, it is reasonable to assume they can think about $(x(t),y(t)) = (t,f(t))$, which conforms to the conventional way of thinking about independent and dependent variables or input-output relationships.

Conceptual Analysis of Rate of Change in Three Dimensions

This section builds on the characterization of quantitative and covariational reasoning above to leverage them in identifying important meanings for rate. Thompson (1994) proposed that rate depends on "coordinated images of respective accumulations of accruals in relation to total accumulations. The coordination is such that the student comes to possess a pre-understanding that the fractional part of any accumulation of accruals of one quantity in relation to its total accumulation is the same as the fraction part of its covariant's accumulation of accruals in relation to its total accumulation" (Thompson, 1994a, p. 237). In other words, given that quantity A and quantity B covaried, if a/b'ths of quantity A has elapsed, then a/b'ths of quantity B has elapsed. Building on a general scheme for rate, Thompson described average rate of change of a quantity as, "if a quantity were to grow in measure at a constant rate of change with respect to a uniformly changing quantity, then we would end up with the same amount of change in the dependent quantity as actually occurred" (Thompson, 1994a, p. 271). As a result, average rate of change relies on a concept of constant rate of change, from which one constructs a concept of instantaneous rate of change while understanding it cannot exist.

Meanings for Rate

A student must have a mature understanding of rate. By this we mean that if a student conceives two quantities A (with measure *a* A-units) and B (with measure *b* B-units) in relation to each other, then to understand that a/b represents the magnitude of quantity A relative to the magnitude of quantity B, the student must think about the statement a/b=*c* as saying that *a* is composed of *c* b-units. In this way, 16/5 = 3.2 says that 16 is composed of 3.2 units of five, or for every five units of B, there are 3.2 units of A. Put yet another way, the student understands that the statement 16/5 = 3.2 says that 16 A-units is 3.2 times as many units as 5 B-units.

Then, to think about instantaneous rate of change as a limit of average rates of change, the student must think about constant rate of change in the following way. Suppose quantity A changes in tandem with quantity B. Quantity A changes at a constant rate with respect to quantity B if, given any change in quantity B and the corresponding change in quantity A, the changes in quantity A and quantity B are in proportional correspondence. In other words, given that quantity A and quantity B covaried, if a/b'ths of quantity A elapses, then a/b'ths of quantity B also elapses. Thus, suppose that a function *f* changes from *f*(a) to *f*(b) as its argument changes from *a* to *b* (Thompson, 1994). The function's output variable (*y*) changes at a constant rate with respect to its input variable (*x*) if whenever a/b'ths of *b-a* has elapsed, then a/b'ths of *f*(b)-*f*(a) has elapsed.

Though instantaneous rate of change exists mathematically, there is no way to talk about instantaneous rate of change within a real world situation without using approximations. Given that one thinks about rate as the constant accumulation of one quantity in terms of the other, we suggest that for a student to think about instantaneous rate of change, he must first understand that if a function *f* changes from *f*(*a*) to *f*(*b*) as its argument changes from *a* to *b*, the function's average rate of change over [*a*,*b*] is that constant rate of change at which another function must change with respect to its input to produce the same change *f* produces over the interval [*a*,*b*] (Figure 2).

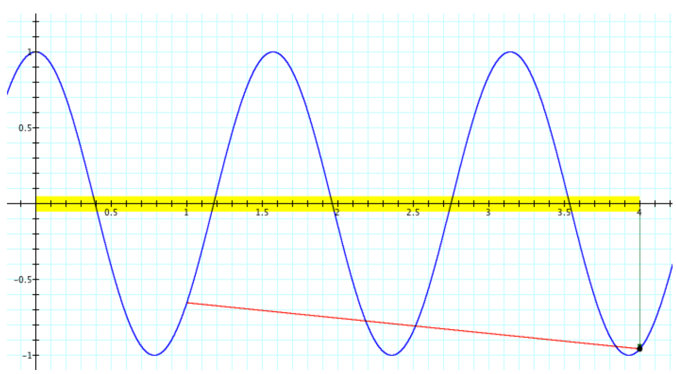


Figure 2. Average rate of change function as constant rate of change

If a student thinks about average rate of change in this way, by attending to the constant rate of change of covarying, accumulating quantities, then instantaneous rate of change is "unreachable". By unreachable, we mean the student cannot viably say that at a specific instant, quantity A is changing at some *exact* rate with respect to quantity B. Instead, this instantaneous rate of change is the result of a finer and finer approximation, generated by considering average rate of change of quantity A with respect to quantity B over smaller and smaller intervals of change for quantity B. The formula for instantaneous rate of change, $f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$, is coherent when the student thinks about f(x+h) – f(x) as quantity A, and h as quantity B.

Rate of change in three dimensions is problematic because it depends on direction in space, yet builds on the foundation of thinking about rate of change of a one variable function as we just described. With a function of one variable, the idea of change is unproblematic—the argument increases or decreases in value and the function increases or decreases accordingly. But with a function of two variables, the idea of change in the argument is ambiguous. The argument can be thought of as a point in a coordinate plane that can change in an infinite number of directions. It does not simply increase or decrease.

Extension to Rate of Change in Space

In traditional notation, the rate of change of f(x,y) in the direction of a vector u = <a,b> can be written as $f_u{}'(x,y)=\lim_{d\to 0}\frac{f(x+ad,y+bd)-f(x,y)}{d}$, which is equivalent to

$$D_u f(x,y)=f_x{}'(x,y)a+f_y{}'(x,y)b$$

It is important to keep in mind that <a, b> is a vector quantity and is the sum of <a, 0> and <0, b>. Thus, the statement $D_u f(x,y)=f_x{}'(x,y)a+f_y{}'(x,y)b$ says that the rate of change of the function f in the direction <a, b> is the vector sum of its rates of change in the directions of its component vectors.

This definition is problematic because it assumes a student understands the necessity and use of a direction vector. It is important that the student thinks about the rate of change of a function as a change in its value relative to a change in its argument, and to understand this in a way that aligns with the way he thinks about the rate of change of a single variable function relative to its argument. To do this, one must imagine picking a point $P_{(x_0,y_0,z_0)}$ on the surface "at" which he or she wants to determine the function's rate of change in a certain direction.

A student must consider what it means for a quantity A to change at a constant rate with respect to quantities B and C. Suppose that the quantities A, B, and C covaried. Quantity A changed at a constant rate with respect to quantities B and C if for any amount of quantity A elapsed (a/b'ths), a/b'ths of quantity B elapsed, and a/b'ths of quantity C elapsed. Extending this to function notation results from thinking about f(x,y) as quantity A, and x and y as quantities B and C. This conception of rate of change necessitates the idea of considering a direction of change, as there are an infinite number of directions in which f(x,y) changes with respect to x and y. The direction actually specifies the relationship between x and y. For example, the rate of change of f(x,y) = 3x + 5y from a point (x_0,y_0) will be constant for any changes in x and y only if Δy is a linear function of Δx.

Instantaneous rate of change can be thought of as an average rate of change over an infinitesimal interval. The average rate of change of a quantity C [f(x,y)] with respect to quantities A (x) and quantity B (y) in a given direction in space can be thought of as the constant rate at which another quantity D would need to change with respect to quantities A and B to produce the same change as quantity C in the same direction that (x,y) changed. This necessitates that quantity D accrues in a constant proportional relationship with quantity A, and simultaneously accrues in a constant proportional relationship with quantity B. These understandings support thinking that as with functions of one variable, an exact rate of change is a construction of an average rate of change between two "points". The points here are (x_0,y_0,f(x_0,y_0)) and (x_1,y_1,f(x_1,y_1)).

Thus, the average rate of change between those two points is the constant rate at which another function g(x,y) would need to change with respect to x and y over the intervals [x_0,y_0] and [x_1,y_1] to produce the same net change as f(x,y) over those same intervals. The function g(x,y) must change at a constant rate with respect to x and a constant rate with respect to y and those constant rates must remain in an invariant proportion, which necessitates x and y accruing in an invariant proportion as well. An "exact" rate of change then, is a result of considering an average rate of change of f(x,y) over an infinitesimally small interval of [x_0,y_0] and [x_1,y_1], where changes in x and y also covary in constant proportion to each other.

The denominator for the rate of change of a two-variable function $f_u'(x,y) = \lim_{h,k\to 0} \frac{f(x+h,y+k)-f(x,y)}{?}$ does not initially make sense because it seems as if h and k vary independently. Thinking about the rate of change of $f(x,y)$ as above supports thinking that any accrual d of either x and y must be made in constant proportion b/a. This proportion a/b actually specifies the direction vector to which many calculus books refer.

Thus, rate of change of f(x,y) with respect to x can be reformulated as $f_u'(x,y) = \lim_{d\to 0} \frac{f(x+ad,y+bd)-f(x,y)}{d}$, where $ad = h$ and $bd = k$ so h must be a/b'ths of k and k must be b/a'ths of h. Then, d can be thought of in the same way as h in the one-variable case, where the student thinks about the derivative as an average rate of change of a function over infinitesimal intervals while realizing that the proportional correspondence between h and k means they have a linear relationship resulting in approaching the point (x_0,y_0) along a line.

Method

Two students, Brian and Neil, participated in a group teaching experiment focused on rate of change. They were taking a first semester calculus course that was grounded in quantitative and covariational reasoning with an explicit focus on understanding rate of change. The teaching experiment sessions contained problems focused on revealing how students think about ideas of rate and their extensions to three dimensions. During these interviews and the teaching experiments, the students were able to use a laptop computer, Graphing Calculator, and had table-sized whiteboards. Each session was videotaped on overhead and side-view cameras, and the screen was recorded using SnapzPro.

Teaching Experiments and Reflexivity

This study built on Steffe and Thompson's (2000) account of a teaching experiment to create a valid and reliable model of student thinking. At the same time, I accounted for my role in what the student said and did and understand that our interpretations of student's behavior and explanations contribute to students' actions, because my actions toward them were predicated on my understandings of what they did. I reflected on how students interpreted my actions and how that interpretation might have played a role in the students' actions. I prepared a set of hypotheses about the student's actions prior to each teaching session, and my hypotheses were based on my working model of their ways of thinking. Immediately after each teaching experiment session, I watched the recording of the session, did a basic transcription, and reflected on how the student's actions within that session necessitated adjustments or changes in the model of their thinking.

Analytical Method

Thompson (2008) described one use of conceptual analysis as to propose models of student thinking that help to make distinctions in students' ways and means of operating, and to explain how these ways of and means of operating persisted or changed during the teaching experiment. The retrospective analyses involved making interpretations and hypotheses about students' thinking that I did not have in the moment of the teaching experiment. As I made interpretations and hypotheses about the students' understandings using retrospective analysis, I used conceptual analysis to continually generate and revise a scheme of meanings that would have made what the students said and did coherent for them.

I documented potential ways of thinking and utterances to suggest ways of thinking that made these behaviors and utterances sensible for the individual student. I used StudioCode to code videos and notes. The coding process centered on theorizing about ways of thinking that explained categories of student behavior by using a combination of open and axial coding (Strauss & Corbin, 1998). The coding of video data served two purposes. It was a way to create objective counts of various coded instances. Second the use of Studiocode to code video instances allowed me to generate videos that contain all instances falling under a particular code, which supported further conceptual analysis of ways of thinking that were allied with the behavior that the code marks.

Key Terms and Definitions

This teaching experiment focused on rate of change, but used terms with which the reader may not be familiar. In the following excerpts, students discuss rate of change in *open* and *closed form* (Weber, Tallman, Byerley, & Thompson, 2012). A function is expressed in closed form when it is defined succinctly in terms of familiar functions, or algebraic operations on familiar functions. The function f defined as $f(x)=\sum_{k=0}^{\infty}\frac{x^k}{k!}$ is in open form, whereas f defined as $f(x)=e^x$ is in closed form. You will see Jesse referring to rate of change functions in open and closed form. The function r_f defined as $r_f(x)=\frac{f(x+h)-f(x)}{h}$, where x varies and h is fixed, is in open form. This function, first used by Tall (1986), gives the average rate of change of f over the interval $[x,x+h]$ for all values of x in f*'s* domain. r_f approximates the derivative of f for suitably small values of h. This simultaneous variation of x and r_f is illustrated by the *sliding calculus triangle*, to which Jesse referred to numerous times in our sessions (Figure 3).

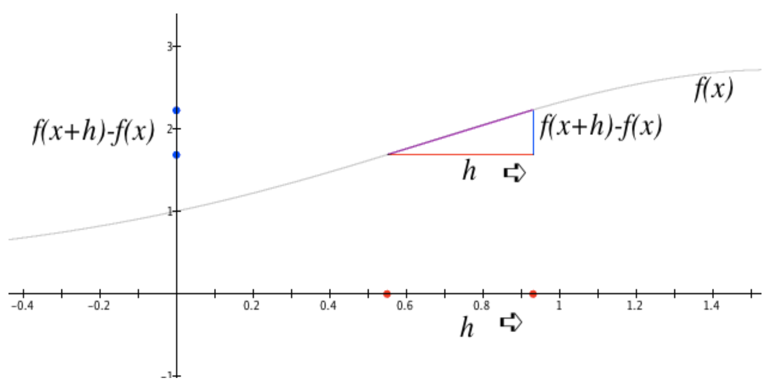


Figure 3. The calculus triangle slides through the domain of the function.

In order to generate outputs for the average rate of change function, one measures the quantity $f(x+h)-f(x)$ in units of h and systematically associates this output value with x, the left endpoint of the interval [x, $x + h$]. Accordingly, a point on the rate of change function can be interpreted as (x, $f(x+h) - f(x)$ units of h). Then, as x varies throughout the domain of the function, this point traces out the rate of change function $r_f(x)$ (Figure 4).

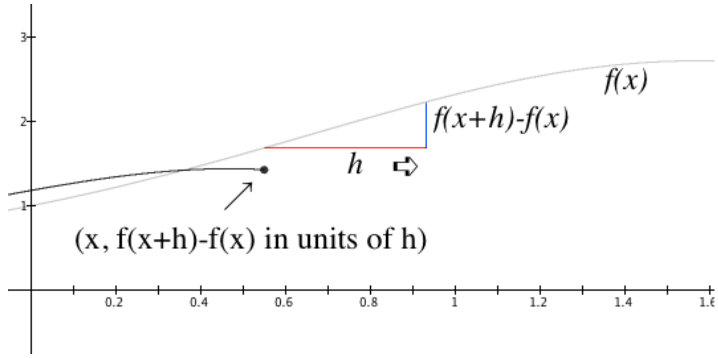


Figure 4. The sliding calculus triangle generates the average rate of change function over intervals of size *h*.

Results: The Meanings Students Developed

In the first part of this paper, I proposed productive meanings for thinking about rate of change that might extend to three dimensions using sophisticated quantitative and covariational reasoning. The results provide excerpts from the two-student teaching experiment to characterize the meanings students possessed for rate and how those meanings affected the development of their understanding of rate in three dimensions.

Rate of Change in Two Dimensions

This excerpt provides insight into the meanings that Brian and Neil had for rate of change in two dimensions. This excerpt occurred at the beginning of the teaching experiment while discussing how to measure and interpret rate for a function in the plane. Neil interpreted rate of change as a measure of steepness of a function, often independent of the quantities he used to measure the rate. In contrast, Brian interpreted the rate of change as how fast one quantity was changing with respect to another. It is crucial to distinguish between how the students interpreted the meaning of the rate and how they measured it. Brian and Neil measured similar quantities to determine the rate, but their interpretation of that calculation suggested that their meaning for rate was different.

Excerpt 1

INT: Alright, so you two seem to be focused on the rate of change of the function graphed here. Can you say more about the rate of change?

Brian: Yeah, well, I guess I was gonna say that you program an input, like where you want to know the rate of change, then use the open form rate of change as the output. But we have talked about using a correspondence point, so the two things you need are x and $r_f(x)$

INT: Okay, and what is $r_f(x)$*?*

Brian: It is the open form rate of change, so for some fixed value of h. So basically it tells you how fast one thing is changing with respect to another over that interval, but tells it to you in a constant rate.

INT: Neil? Did you want to add anything?

Neil: Yeah, I agree with the constant rate of change, straight line kind of thing, but know that it tells me the slope of a graph, or how steep a function really is. A rate tells me a slope of a graph, whether or not we have the graph of the function.

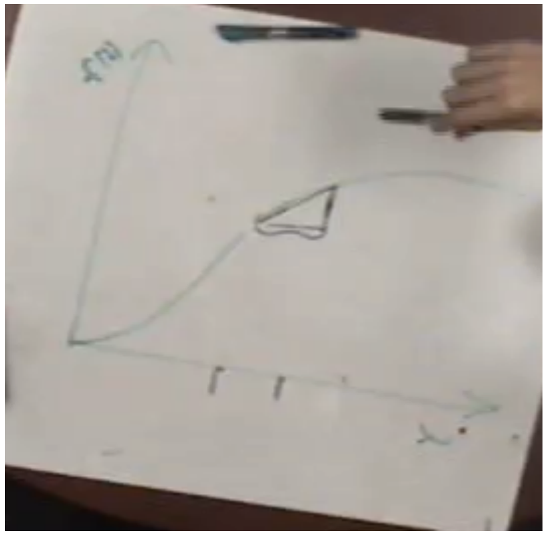


Figure 5: Brian and Neil's illustration of how to measure rate of change on a graph.

Extension to Three Dimensions

A major theme of the teaching episodes was Brian and Neil's developing understanding of the relationship between average and instantaneous rate of change. Brian and Neil developed an interpretation of instantaneous rate as an "in the moment" measure of steepness (Neil) and how fast quantities were changing (Brian). In the moment was their mnemonic intended to distinguish instantaneous rate from average rate of change. After Brian and Neil had working experience in constructing and interpreting graphs of two-variable functions, they discussed how to measure and interpret rate of change at a point in space.

Excerpt 2

INT: How would you think about rate of change at a point in space?

Brian: My first though is that it has multiple rates of change, kind of like sitting on a hill, depending where you look, the steepness, slope at that point can be different.

Neil: I agree with that, I thought about the kind of example too, or just sitting on the surface we swept out, maybe we can use the z-x and z-y rates of change?

Brian: Yeah! Umm, let's see though, if we want to make a rate of change function and then graph it in space, we need to figure out a way to program it, and do a sweeping out.

Neil: What about just plugging in x, y and the rate of change? Oh, I guess we don't know the rate of change yet, so my point thing wouldn't work.

Brian: Alright, for z-x, it is kind of like, okay let's back up, let's say we are at a point (a,b,c) in space. Then for z-x, we fix y at b, then do the normal rate of change except it has to be two-variables.

Neil: Yeah, that makes sense, so we need an h, maybe like f(x+h,b)-f(x,y), then divided by h, and then for z-y, we just say y+h and fix x at a?

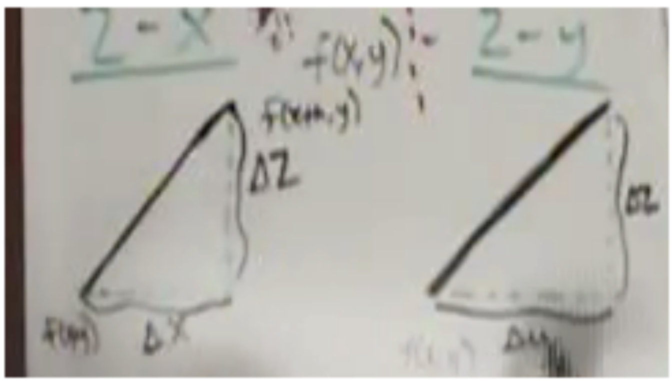
Brian: Yeah, let's go with that.

INT: Okay, so where do you want to go next then, what is your plan?

Brian: We need the two rates of change to make an overall rate of change function, then we can graph it by doing the sweeping out I think.

Neil: I'd rather just draw the two calculus triangles that I am imagining each in a perspective.

Brian's description of rate of change indicated he was thinking about rate of change in a direction "at" a point on the surface of the function's graph. Brian's suggestion of considering multiple rates of change from a perspective (e.g. z-x) led to Neil's sketching of perspective dependent calculus triangles (Figure 6). Their descriptions of multiple rates of change, as well as their determination of an average rate of change function for the z-x and z-y perspectives, indicated they were imagining rate of change occurring in at least two directions. Their sketch of open form rate of change functions for both the z-x and z-y perspectives indicated they agreed on a method to measure the rate. However, they both remained resolute in their interpretation of the measurement.



$$r_{f_x}(x,y) = \frac{f(x+h_1,y)-f(x,y)}{h_1} \quad (z\text{-}x)$$

$$r_{f_y}(x,y) = \frac{f(x,y+h_2)-f(x,y)}{h_2} \quad (z\text{-}y)$$

Figure 6. Calculus triangles from the z-x and z-y perspectives.

Direction in Space

Their measurement of partial rates of changes led to a debate spurred by Neil's claim that "the" rate of change could be determined at a point on a surface in space. Brian and Neil understood that they could measure two rates, but struggled to combine these rates to produce a "total rate of change". Total rate of change was their description of a point on a surface that "possessed" a single rate. Their debate led to their understanding the utility of considering direction in space. In the following excerpt, I asked Neil and Brian to expand on their description of their perspective dependent calculus triangles, in particular their use of h_1 and h_2 in the denominators of the open form rate of change functions. I anticipated that thinking about the relationship of h_1 and h_2 would be critical to their creating a need for considering rate of change in a direction and would help them resolve their debate about a "total rate of change".

Excerpt 3

INT: *So I noticed that you guys constructed your calculus triangles in each perspective and labeled the h's as different. Can you say more?*

Neil: *Yeah, well basically the h's are independent, so they don't have to be the same, but I guess they could be.*

Brian: *I was thinking about this more last night, and had a sort of moment. We talked earlier about being on a hill, or walking on a function, for example, and the rate of change depended on the direction you were facing, I think the same thing applies here.*

INT: *What do you mean by a direction?*

Brian: *Okay, so let's say we head from a point, or we went directly Northeast, and imagine we are doing this from overhead, it was looking at the x-y perspective that made me think of this.*

INT: *Okay, so if you were heading Northeast, what's the significance of that?*

Brian: *Then we know how the two h's are related to each other, because most of the time you don't go in a direction of just north, west, east, whatever, you head in some combination.*

Neil: *Okay, not sure I am following, but basically if we head Northeast, the change in x and change in y would be equal, a different direction, change in x could be twice as big as change in y, which are like the h's right?*

Brian: *Yeah, let me make an illustration here. Then we can just call the numerator a change in z,* $f(x+h_1, y+h_2)-f(x,y)$ *either h value that we want to.*

Brian introduced direction as a way to account for all possible rates of change in space. I believed Brian had an image of the relationship between h_1 and h_2 to define a direction. Brian's key insight was that *any* direction was more general than considering only z-x and z-y perspectives. Neil and Brian continued to develop a two-variable open form rate of change function, and they agreed that the numerator represented a change in the output, represented by $f(\text{x}+h_1, \text{y}+h_2)-f(\text{x},\text{y})$, where either h_1 or h_2 was written in terms of the other *h*-value. However, they questioned how they have a single denominator that represented a change in *x* and a change in *y*. Even though they saw that h_1 and h_2 depended on each other, that dependence did not immediately resolve their issue of what change to represent in the denominator. They appeared to think picking h_1 or h_2 would happen at the exclusion of the other.

Excerpt 4

INT: *So, what is your denominator in the function, your conjecture?*

Brian: *I was thinking either h-value, whichever you have in the numerator.*

Neil: *But doesn't that kind of just delete it, it goes away?*

Brian: *No, you just define it in terms of* h_1*, if you are talking about* h_2.

INT: *So if one h is in the denominator, and we make that value small, what happens to the other h value?*

Brian: *Well, oh yeah, it becomes small as well, it's not like the h's have the same value, but they can end up getting so small it doesn't matter.*

Neil: *Ah, I see what you mean. So because we have an equation to relate the two h's, then when one gets really small, the other one has to get smaller too? I was thinking multiply in the denominator I guess, but that doesn't need to happen, you just need one of them to become small because they are related, the h's I mean, so then both do.*

Brian: *So, I guess that's our overall rate function for two variables. Now we can sort of program the points to do the graph.*

Brian's insight that the changes in x and y became smaller in tandem allowed him to conjecture that using a single parameter in the denominator was acceptable. Neil appeared to focus on deleting one of the parameters, but Brian's insight allowed him to think about the equation they had specified between h_1 and h_2. By imagining progressively smaller values for h_1, he found that h_2 became smaller as well given the proportional relationship specified by choosing a direction in space (see Figure 7). These insights allowed them to construct an average rate of change function, or open rate of change function in their terms (Figure 7).

$$r_f(x,y) = f(x+h_1, y+h_1 d) - f(x,y) \Big/ h_1$$

direction = d d= fixed $\frac{\Delta y}{\Delta x}$ $h_1 = \Delta x$ $h_2 = \Delta y$

Figure 7. Brian and Neil's two-variable open form rate of change function.

Brian and Neil made a number of sophisticated observations in their construction of the open form rate of change function. Throughout the teaching experiment they returned to their meaning for rate of change that they had identified in two dimensions. I asked them to reflect on their description of rate of change as steepness (Neil) and rate of change as a measurement of covariation (Brian) to understand in what ways they thought their meanings had changed.

Excerpt 5

INT: *So, now we have had a lot of time to talk about rate of change, I wanted to go back to where you talked about interpreting the value. So if you determined that the rate of change at a point in space was 3.2, how would you interpret that value?*

Neil: *Well, again, the 3.2 represents a slope of a graph, this time it is a little more sophisticated slope than we had before, but still steepness like on a hill. The same as in two dimensions, just using a different way to find the rise over the run than before.*

Brian: *I can see where Neil is coming from, but I think of it more as a consistent thing with two dimensions, where it tells you how fast one thing is changing with respect to another, except the another in this case is two other quantities with a fixed change between the two.*

INT: *Fixed change?*

Brian: *Yeah, like one changes twice as much as the other.*

Neil: *See, I am not sure about this how fast business. We are describing a graph still, how fast does not make as much sense to me because I don't see the speed idea.*

Discussion

Measurement and Interpretation of Rate

The discussion in Excerpt 5 brings out a distinction between how Brian and Neil thought about 1) measuring the rate of change and 2) the meaning they had for that measurement. The previous excerpts showed that they worked in tandem to produce an open form rate of change function that used direction in space in a sophisticated way. They also made insights about proportional relationships between changes in quantities as a result of specifying a direction. However, it appeared that the imagined measuring different things. Brian imagined measuring the changes in quantities to determine how fast they were changing, understanding that a graph was the representation of how those quantities changed in tandem. Neil imagined measuring changes because he wanted to determine a rise and run in the process of calculating steepness. The differences in their meanings and the sophisticated advances they were able to make while using these meanings for rate of change have implications for how we think about studying meaning and students' mathematical knowledge.

Brian and Neil's development of an open form rate of change function required them to make a number of insights about direction in space, proportionality of changes in quantities while leveraging desire to represent the rate using a quotient. It was most surprising that their meaning for rate of change changed little, even as it allowed them to make a sophisticated generalization about how to measure rate of change in three dimensions. They participated in a number of teaching episodes specifically designed to engender ways of thinking about constant, average, and instantaneous rate as I described in the conceptual analysis, yet their meanings were robust enough to help them make sense of all of those situations. It is possible that their meanings shifted in the moment, but any shift was temporary. The persistence of their understanding of rate suggests the students possessed a somewhat complicated scheme that allowed them to reason about many situations and tasks while retaining their meaning. In Neil's case, he was able to make advances in measuring rate while thinking about it solely in a graphical context.

Generalization and Abstraction

Neil and Brian were able to make sophisticated generalizations in their ability to measure rate of change, yet their meanings did not shift. A natural question that emerges from this is: what did the students generalize? I believe that they generalized both the structure of a quotient to measure rate and their meaning for it. However, only in Brian's case was this generalization the one intended. These results suggest that it is important to understand how students' meanings for mathematical ideas affect their ability to "generalize" those meanings. This focus on generalization is not new, but it has not yet been applied studying the transition students make from single to multivariable calculus. For example, Ellis (2007a, 2007b) proposed frameworks for considering the generalizing actions that student undertake. Simon, Tzur, Heinz, and Kinzel (2004) proposed that reflective abstraction is at the heart of students achieving more advanced ways of thinking based on their current understanding and ways of thinking. Together, Ellis and Simon have described frameworks for considering the ways of thinking and the actions necessary to support development of more advanced ways of thinking. At the same time, Simon and Tzur (2004) described a hypothetical learning trajectory as a way to both generate and test hypotheses about the development of student thinking and the role of tasks in that development. The results of this paper suggest that is critically important to understand the role of students' meaning in the generalizations they make in mathematics, and one avenue to that may be construction of hypothetical learning trajectories that are sensitive to meaning and generalization.

Effective construction of these hypothetical-learning trajectories require not only a single instantiation of a study, but also an iterative cycle of both hypothesis generation and testing. This study can be considered a starting point on which future iterations can build. The two-change problem and the results I have presented here suggest that the transition from single to multivariable calculus is a fruitful place to begin this exploration.

References

Carlson, M. P., 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.

Ellis, A. (2007a). Connections between generalizing and justifying: Students' reasoning with linear relationships. *Journal for Research in Mathematics Education, 38*(3), 194-229.

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

Saldanha, L., & Thompson, P. (1998). Re-thinking co-variation from a quantitative perspective: Simultaneous continuous variation. In S. B. Berensah, K. R. Dawkings, M. Blanton, W. N. Coulombe, J. Kolb, K. Norwood & L. Stiff (Eds.), *Proceedings of the Twenieth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 298-303). Columbus, OH: ERIC Clearninghouse for Science, Mathematics, and Environmental Education.

Simon, M., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning, 6*(2), 91-104.

Simon, M., Tzur, R., Heinz, K., & Kinzel, M. (2004). Explicating a mechanism for conceptual learning: Elaborating the construct of reflective abstraction. *Journal for Research in Mathematics Education, 35*(5), 305-329.

Steffe, L. P., & Thompson, P. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Strauss, A. L., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

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

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

Thompson, P. (in press). In the absence of meaning. In K. Leatham (Ed.), *Vital directions for research in mathematics education*. New York: Springer.

Weber, E., Tallman, M., Byerley, C., & Thompson, P. (2012). Introducing derivative through the calculus triangle. *Mathematics Teacher, 106*(4), 274-278.