PROCEEDINGS OF THE 14TH ANNUAL CONFERENCE ON RESEARCH IN UNDERGRADUATE MATHEMATICS EDUCATION

EDITORS
STACY BROWN
SEAN LARSEN
KAREN MARRONGELLE
MICHAEL OEHRTMAN

PORTLAND, OREGON
FEBRUARY 24 – FEBRUARY 27, 2011

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

CITATION: In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education*, Vol. #, pg #-#. Portland, Oregon.

CONFERENCE PROGRAM COMMITTEE

STACY BROWN, PROGRAM COMMITTEE CHAIRPERSON

KAREN MARRONGELLE , LOCAL ORGANIZER

JASON BELNAP

JENNIFER CHRISTIAN-SMITH

NICOLE ENGELKE

TIMOTHY FUKAWA-CONNELLY

HOPE GERSON

KAREN ALLEN KEENE

JESSICA KNAPP

SEAN LARSEN

MICHAEL OEHRTMAN

ALLISON TONEY

JOSEPH WAGNER

AARON WEINBERG

WITH MUCH APPRECIATION WE THANK THE CONFERENCE REPORT AND PROCEEDINGS REVIEWERS

JASON BELNAP

STACY BROWN

TODD CADWALLADEROLSKER

JENNIFER CHRISTIAN-SMITH

PAUL DAWKINS

NICOLE ENGELKE

TIMOTHY FUKAWA-CONNELLY

HOPE GERSON

ESTRELLA JOHNSON

KAREN ALLEN KEENE

JESSICA KNAPP

YVONNE LAI

SEAN LARSEN

ELISE LOCKWOOD

MICHAEL OEHRTMAN

KEVIN MOORE

KYEONG HAH ROH

ANNIE SELDEN

CRAIG SWINYARD

ALLISON TONEY

HORTENSIA SOTO-JOHNSON

JOSEPH WAGNER

AARON WEINBERG

CONFERENCE PARTICIPANT LIST

Veda Abu-Bakare
Aditya 'Adi' Adiredja
Aina K. Appova
Homer Austin
Jathan Austin
Anna Bargagliotti
Joanna Bartlo
Mary Beisiegel
Jason Belnap
Kavita Bhatia
Tyler Blake
Tim Boester
William Bond
Jim Brandt
Stacy Brown
Martha Byrne
Todd CadwalladerOlsker
Mindy Capaldi
Laurie Cavey
Sergio Celis
Danielle Champney
Simin Chavoshi
Sean Chorney
Warren Code
Derron Coles
Darcy Conant
Sandy Cooper
Beth Cory
Paul C Dawkins
Louis Deaett
Jason Dolor
Joseph Ediger
Jess Ellis
Rob Ely
Nicole Engelke
Sarah Enoch
Jodi Fasteen
Jill Faudree
Leanna Ferguson
Brian Fisher
Tim Fukawa-Connelly
Evan Fuller
Gillian Galle
Tyler Gaspich
Tyler Gaspich
Hope Gerson
Sylvia Giroux
David Glassmeyer
Jim Gleason
Shiva Gol Tabaghi
Zahra Gooya
Mairead Greene
Todd Grundmeier
Beste Gucler
Taras Gula
Aviva Halani
Catherine Hart-Weber
Mark Haugan

CONFERENCE PARTICIPANT LIST

Shandy Hauk
Krista Heim
Francesca Henderson
Kate Horton
Aladar Horvath
Estrella Johnson
Gulden Karakok
Karen Allen Keene
Rachael Kenney
Minsu Kim
Margaret Kinzel
Jessica Knapp
Libby Knott
Marina Kogan
Janet Kolodner
Dave Kung
OhHoon Kwon
Yvonne Lai
Elaine Lande
Bryan Lane
Sean Larsen
Christy Larson
Christine Latulippe
Sandra Laursen
Elise Lockwood
Tom Lougheed
Tim Lucas
Dann Mallet
Ami Mamolo
Aaron Marmorsdtein
Karen Marrongelle
Sarah Marsh
Jason Martin
John Mayer
Carolyn McCaffrey
David Meel
Pablo Mejia-Ramos
Kate Melhuish
Kelly Mercer
Vilma Mesa
David Miller
Briana Mills
Melissa Mills
Susanna Molitoris Miller
Kevin C. Moore
Ricardo Nemirovsky
Kristin Noblet
Jennifer Noll
Michael Oehrtman
Jeanette Palmiter
Eric Pandiscio
Frieda Parker
Valerie Peterson
Costanza Piccolo
Kirthi Premadasa
Jeffrey Rabin
Sonya Redmond
Kathryn Rhoads

CONFERENCE PARTICIPANT LIST

Lisa Rice
Kyeong Hah Roh
Rebecca Ross
Sarah Rozner
Aron Samkoff
Jason Samuels
David Meel
Milos Savic
Vicki Sealey
Annie Selden
John Selden
J. Michael Shaughnessy
Mary Shepherd
Olga Shipulina
Paula Shorter
Carole Simard
Ann Sitomer
Michael Smith
Hortensia Soto-Johnson
Natasha Speer
Lyn Stallings
Stephen Strand
April Strom
Heejoo Suh
George F. Sweeney
Craig Swinyard
Jen Szydlik
Michael Tallman
Gail Tang
Eva Thanheiser
Matt Thomas
John Thompson
Anna Titova
Abdessamad Tridane
Maria Trigueros
David Tsay
Carla van de Sande
Margarita Vidrio
Sasha Wang
Aaron Wangberg
Laura Watkins
Megan Wawro
John Weber
Keith Weber
Thomas Wemyss
Ian Whitacre
Tim Whittemore
Mark Yanotta
Nissa Yestness
Michelle Zandieh
Dov Zazkis
Rina Zazkis

FOREWORD

The research reports and proceedings papers in these volumes were presented at the 14th Annual Conference on Research in Undergraduate Mathematics Education, which took place in Portland, Oregon from February 24 to February 27, 2011.

Volumes 1 and 2, the RUME Conference Proceedings, include conference papers that underwent a rigorous review by two or more reviewers. These papers represent current important work in the field of undergraduate mathematics education and are elaborations of the RUME conference reports.

Volume 1 begins with the winner of the best paper award, an honor bestowed upon papers that make a substantial contribution to the field in terms of raising new questions or gaining insights into existing research programs.

Volume 3, the RUME Conference Reports, includes the Contributed 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.

Volume 4, the RUME Conference Reports, includes the Preliminary Research Reports that were presented at the conference and that underwent a rigorous review by at least three reviewers prior to the conference. Preliminary Research Reports discuss ongoing and exploratory research studies of undergraduate mathematics education. To foster growth in our community, during the conference significant discussion time followed each presentation to allow for feedback and suggestions for future directions for the research.

We wish to acknowledge the conference program committee and reviewers, for their substantial contributions and our institutions, for their support.

Sincerely,

Stacy Brown,
RUME Organizational Director & Conference Chairperson

Sean Larsen,
RUME Program Chair

Karen Marrongelle
RUME Co-coordinator & Conference Local Organizer

Michael Oehrtman
RUME Coordinator Elect

VOLUME 1

CONFERENCE PROCEEDINGS PAPERS

VOLUME 1
TABLE OF CONTENTS

ANALYZING THE TEACHING OF ADVANCED MATHEMATICS COURSES VIA THE ENACTED EXAMPLE SPACE

Tim Fukawa-Connelly, Charlene Newton and Mariah Shrey
The University of New Hampshire
Tim.fc@unh.edu

In advanced undergraduate mathematics, students are expected to make sense of abstract definitions of mathematical concepts, create conjectures about those concepts, write proofs and exhibit counterexamples of these abstract concepts. In all of these actions, students may draw upon a rich store of examples in order to make meaningful progress. We have drawn on the concept of an example space (Watson & Mason, 2008) for a particular concept. We adapted it and defined the concepts of example neighborhood, methods of example construction, and the functions of examples to create a methodology for studying the teaching of proof-based courses. We demonstrate our method via a case study from an undergraduate abstract algebra course.

Keywords: teaching, examples, definitions, proof, abstract algebra

Introduction

Studying teaching is by nature a difficult process. For this reason, “very little empirical research has yet described and analyzed the practices of teachers of mathematics” (Speer et al., 2010, p. 99) at the undergraduate level despite repeated suggestions for this type of study (Harel & Sowder, 2007; Harel & Fuller, 2009). That is, “researchers’ questions, methods, and analyses have not generally targeted what teachers say, do, and think about collegiate classrooms in an extensive or detailed way” (Speer et al., 2010, p. 105). Although it is important, in and of itself, to better understand the reality of collegiate mathematics classes, it is also essential to develop research-based descriptions of traditional undergraduate classes to support and explain studies of student’s mathematical proficiency.

Within the context of the advanced undergraduate mathematics classes one area of increasing emphasis is on the use of examples. First, research indicates that “exemplification is a critical feature in all kinds of teaching, with all kinds of mathematical knowledge as an aim.” (Bills & Watson, 2008, p. 77). This has resulted in the study of example usage among students. For example, Alcock and Inglis (2008) examined doctoral students’ use of examples in evaluating the truth-value of claims as a way to understand what students might do. Dahlberg and Housman (1997) found that students who generated their own examples were more likely to develop initial understandings of concepts, and Mason and Watson (2008) described ways to make use of the range of possible variation for pedagogical purposes, to name but a few. Yet, these studies are on students’ uses of examples and there is no corresponding study of the teaching of proof-based classes and instructor’s teaching with examples.

In the following article we discuss the current state of research on examples as tools for learning mathematics and draw implications for teaching. We then use those implications to outline a methodology for analyzing undergraduate mathematics teaching through the lens of examples. Finally, we show an example of using the method by analyzing the instruction of an introductory abstract algebra course In the case study that follows we describe the teaching of the class and then analyze the teaching via the lens of the enacted example space.

Literature Review

1.1 Studying teaching

At the undergraduate level, little is known about the processes of teaching and learning of proof-based mathematics courses. Mejia-Ramos and Inglis (2009) conducted a literature search and found only 102 educational research papers that studied undergraduate students' experiences writing, reading and understanding proof. Of those 102 papers, Mejia-Ramos and Inglis found no papers describing how students understand instructor presentation of proof. There is only one study specifically designed to investigate how instructors use examples to support their teaching of proof-based classes. There are four studies that give some description of how instructors use of examples in teaching.

The first was specifically drawing on student's use of examples as a way of understanding classroom activities. Larsen and Zandieh (2008) adapted Lakatos' (1976) lens of proofs and refutations from the way that mathematics might be created to describe how students might recreate mathematical ideas. They first described how examples are used to create and modify definitions, which they called "monster barring." Second, they described "exception barring" as a means of creating or modifying a conjecture. Finally, they described proof analysis as a further way to refine a conjecture and generate new concepts. At each step they showed evidence of students using examples in the specified way drawn from the context of an inquiry-oriented abstract algebra class.

Weber (2004) analyzed the teaching of a real analysis class with a focus on how the instructor presented proofs. He identified three basic styles of teaching proof: *logico-structural*, *procedural*, and *semantic*. The first two styles specifically avoided any use of examples or diagrams. The last, the *semantic* teaching style, is characterized by the instructor's use of intuitive descriptions of concepts and relationships. In each case, the instructor's teaching style was chosen to help students acquire some specific type of knowledge or ability needed to construct proofs (as well as to demonstrate a large number of similar proofs). The semantic style was meant to help the students learn how an understanding of the concepts of mathematics, when supported by procedural fluency with the logic and language of mathematics, would support proof writing. An example of a proof-presented in this style that includes example usage would be one where the instructor divides the board into halves and presents the proof on one half while showing each part in the case of a specific example on the other half. The instructor specifically said that the goal was for the students to "have rich imagery that they could associate with the concepts being taught" (Weber, 2004, pp. 126-127).

While the instructor in Weber's paper claimed that he only used semantic style proofs at the end of the course, many mathematicians claim that examples are important both to their own work and to their teaching. Alcock (2009) and Weber (2010) each interviewed a collection of mathematicians about their teaching practices that support the learning of proof. The two groups of mathematicians interviewed described many similar practices and both groups emphasized the importance of teaching students to instantiate by assigning students problems that require example generation (Alcock, 2009). Similarly, eight of the nine mathematicians in Weber's (2010) study said that they normally accompany a proof with an example or draw a diagram that is intended to support students' understanding of the proof. Given that examples are believed to support students' developing understanding of concepts and proof-writing abilities, it is important to better understand why examples are important in mathematics and the learning of advanced mathematics.

1.2 What are examples, and, what will we study?

This paper draws upon Watson and Mason's definition of an example as "any mathematical object from which it is expected to generalize" (2005, p. 3). All further discussion of the role and importance of examples will rely upon an important pedagogical distinction between examples of a concept (such as group, ring, field) and examples of a process (such as using the Euclidean Algorithm to determine the GCD). While both can be understood as mathematical objects from which to generalize, the paper that follows will draw upon examples of concepts in mathematics teaching, though we assert a similar methodology could be used to study examples of processes. Examples of concepts are uniquely powerful in both mathematics and the teaching of mathematics and, as a result, there are numerous reasons to study the use of examples in classes.

1.3 Why examples are so important in advanced mathematics and mathematics teaching

Mathematicians and students use collections of examples as references to develop intuition and as a means to generate, test, and refine conjectures (Alcock & Inglis, 2008; Michener, 1978). When a mathematician comes upon or creates a conjecture, hypothesis, or theorem that is not obviously true, Courant (1981) claimed that the mathematician's first reaction was to call upon an example so as to think about the general through a particular case. Goldenberg and Mason (2008) stated that there is little difference between examples and counterexamples when testing and refining conjectures; the difference lies in what the reader attends to. In other words, an example that demonstrates the truth of the theorem can be considered a non-example of an incorrect version of the theorem, and a counterexample to a proposed theorem allows for reformulation of an incorrect claim into a correct claim. In either case, the purpose of the example is to provide a more familiar and concrete means to explore ideas and to check the conditions of and evaluate constraints in theorem formulation.

These claims for the mathematical importance of examples have also shaped the way that examples are used in teaching. Examples are often used to introduce and motivate topics in class. In particular, examples that introduce concepts give individuals a concrete and potentially familiar means to explore and understand the constraints and affordances of a definition. Dienes (1963) argued that mathematics learners require at least three examples of a concept in order to develop understanding. That is, examples give the student a collection of familiar objects from which they can develop abstractions that are eventually reified into concepts (Sfard, 1994). Similar to Goldberg and Mason's (2008) claims about non-examples helping individuals to understand the conditions of a theorem, Zazkis and Leiken (2008) refer to pertinent non-examples in the situation of testing conjectures for truth-value because non-examples also help mathematics learners make sense of particular aspects of definitions by asking, "Why does this not satisfy the definition?" Thus, examples and pertinent non-examples give learners more opportunities to gain a complete understanding of each clause of a given definition, both of what that definition allows and of what it prohibits in terms of structures. Moreover, as Goldenberg and Mason posited, it is by exploring examples that "learners encounter nuances of meaning, variation in parameters and other aspects that can change" (2008, p. 184).

While examples and non-examples that students encounter from the curriculum are important in helping them to develop an understanding of important concepts, Zazkis and Leiken (2008) also describe the importance of learner-generated examples. They claim that learner-generated examples are especially for instructors. Instructors can evaluate their students' understanding of the definition by asking them to construct examples. Therefore, they can be used to diagnose problematic aspects of the student's understanding. Thus, examples serve as a type of formative assessment that allows faculty to propose structures (either pertinent non-examples or examples) to better help the students develop their understanding of the concept.

In summary, examples can be used in mathematics and mathematics teaching for a number of different functions. Lakatos (1976) described the uses of examples as articulating and refining definitions and conjectures, that is, as objects from which to generalize. Mathematicians use examples to attempt to check the validity of conjectures and as clues to the proofs of theorems. In teaching, examples provide a way for students to attach meaning to definitions (Goldenberg & Mason, 2008). Similarly, they can also help develop understanding of proofs for appropriate theorems (Alcock, 2009; Weber, 2010). We believe that the range of functions that students are exposed to is important in their understanding of the field of mathematics. Essentially, students should see a wide range of examples functioning in different ways: exemplifying definitions, helping to generate and test conjectures and as providing clues to proofs.

Mason and Watson (2008) have taken the mathematical and pedagogical purposes of examples and created a description of a construct that mathematics users might have. They described the *example space* as a construct that allows mathematics users to develop an appropriate sense of the concept being exemplified and to learn to adequately perform all of these other functions of examples.

1.4 Example spaces: Students' range of thought, knowing what can vary, knowing what must stay the same

An example space is the "experience of having come to mind one or more classes of mathematics objects together with construction methods and associations" (Goldenberg & Mason, 2008, p. 189). It may include relatively frequently accessed members of the class, less accessed members, and new members (via construction methods). Two features of example spaces that Mason and Watson (2008) called important: what aspects of the examples the learner realizes can be varied, and what range of variation the learner believes is appropriate.

Goldenberg and Mason (2008) draw upon variation theory and claim that learners need experience with many examples. The first examples that learners experience are particularly important as they are often the ones that students most closely link with the concept, and, as a result, should be very carefully chosen (Zodik & Zaslavsky, 2008). Students may modify their understanding of the definition of the concept based upon their image of the concept (Vinner, 1991). As a result, when students experience an early example or set of examples that they adopt as their concept image, it can shape their understanding of the concept itself. Thus, the early examples need to help students develop a complete concept image with all appropriate complexity. Vinner suggests that "only non-routine problems, in which incomplete concept images might be misleading, can encourage people to" develop more appropriate understandings (1991, p. 73).

Golderberg and Mason (2008) further the claim that the learners need to work with a carefully chosen set of examples closely in time. They believe that this will allow learners to "locate dimensions of possible variation… [and] discern which aspects can vary and which are structural" (Goldenberg & Mason, 2008, p. 186). Dienes (1963) argued students should encounter sets of examples that are narrowly constructed so that examples within each set vary along only a very limited number of dimensions. Thus, we believe that students should work with multiple sets of examples, each with variations along different dimensions so that they are able to apprehend both what aspects of an example can be varied and the range of possible variation for each aspect.

If we believe the arguments put forth about the importance of showing students a range of examples with different variation in carefully constructed ways, it follows that examining the collection of examples of a certain topic that students are exposed to and the order in which they encounter them could allow insight into the types of understandings that students will develop in terms of what can be varied and the range of possible variation. Thus, the range and organization of examples is directly linked to possible student learning. Of course, mathematics learners can construct appropriate or inappropriate beliefs from the collections of examples.

For example, when learning the definition of a group, the possible aspects of a group include characteristics of the underlying set, the group itself, and of the behavior of specific elements. In particular, we would want a learner to believe that the size of the set of a group is immaterial to the definition. In terms of the range of possible variation, we would want the learner to recognize that the cardinality could vary from 1 to groups with infinite cardinality.

A final important feature of an example space, in the way that Goldenberg and Mason (2008) have defined it, is that it purposefully includes construction methods and associations, such as links to important theorems and relations to other constructs. These links allow mathematicians and mathematics learners to create new examples that meet specific criteria of theorems and to discover which classes of objects are most relevant in particular situations.

Our Theory of Assessing the Enacted Example Space

We have drawn on the work in Section 1.4 to articulate a method for examining the enacted example space that uses three filters to describe the set of examples. We call these filters: (1) *example neighborhood*, (2) *example construction*, and (3) *the function of the example*.

First, we define the *example neighborhood* as the entire collection of examples that the students are exposed to during the course of their studies of a particular construct. These may be concrete examples or relevant non-examples of a given concept. A typical *example neighborhood* is the sequence of examples given to support the definition of a concept. We call this type of neighborhood a "definition-example" neighborhood. Another type of neighborhood arises in student-worked problems initiated with a stem such as, "Determine whether the structures below are examples of a _________."

We analyze how the examples in the example neighborhood are organized on four levels: (1) what is the first example given, (2) what examples are near the concept temporally, (3) what is the range of permissible variation that students experience, and (4) what variation constraint do students experience. We pay particular attention to the first few examples, as instructors believe they are often the ones that students most closely link with the concept (Zodik & Zaslavsky, 2008). Examples have less immediate relation to the central example(s) when the further they are removed either temporally or conceptually, depending on the number of aspects that vary from the central example(s). In terms of the range of permissible variation we assess what examples the students are given access to and the ranges of variation that those encompass. Similarly, we look at the set of non-examples that the students have access to and how those non-examples limit the concept in question. In assessing the variation constraint, we adopt Dienes (1963) argument that students should see examples that vary only in a constrained manner so that they can determine what is structural and what is allowed to vary, as well as apprehend the range of permissible variation. Then, he argued, they should see other examples that vary along a different dimension. As a result, we argue that early examples that vary along too many dimensions may actually lower the potential value for student learning. Similarly, a collection of examples that fails to support student construction of critical aspects of the concept will also lead to lessened possible student learning.

Secondly, we examine *example construction* to support a particular concept. Example construction focuses on the range of possible variation included in the neighborhood of a particular example space. The analysis of example construction focuses on how examples are created and the tools for creating additional examples. Example construction also allows mapping from concrete examples to a broad description of the example space that students may be able to populate themselves. In this way, the example space explicitly includes both examples and the means of construction (Goldenberg & Mason, 2008).

Third, we describe the *function of the examples* used in the classroom. We mean this in two different ways. First, we examine which examples are called upon most often. Frequently used examples may obtain "ready access" status for students (linked to Vinner's (1991) concept of *evoked concept image*). The frequency of use gives us a means to assess or predict the student's perception of the relative importance of each example and provides a way to predict which examples can most readily function as an example for the students. Secondly, we describe what the example was being used to do, that is, the mathematical intent of the example. Some possible functions of examples are exemplifying a definition, creating or refining of a definition (c.f., Larsen & Zandieh, 2008), articulating or exploring a conjecture, as well as illustrating a proof. We assess examples separately using each filter, and then read them together to analyze the example space. Taken together, these three rounds of assessment of the example space contribute to our proposed method of assessing the mathematical quality of instruction at the advanced undergraduate level.

We do believe that there is a possible balance that may be achieved by professors in giving the students the ability to populate their example spaces. We can imagine an instructor that directly gives the students access to a wide range of examples but few methods of construction. This range of examples would still give the students the possibility of developing a rich example space. At the other extreme, the instructor may give the students access to a relatively few examples, but a rich set of tools for example construction and a set of tasks that requires and supports the students in developing their own examples. We do not hold either of these extremes to be normative, but do believe that it is important that the students have the opportunity to develop a well-defined example space.

While it is true that the depth with which students engage the examples matters for their learning (Vinner, 1991), it is impossible to make any low-inference judgments about this from classroom observation. Students may be engaging with the material as it is presented in a lecture but, because they are not giving outward sign it is impossible to make inferences. Similarly, in an inquiry-based class, while students may be observed engaged in active discussions in small group, unless each group is monitored it is impossible to make judgments about their level of engagement. Other students presenting at the board is similar to the situation found in a lecture. Finally, even with suggested or required homework problems there is no way of evaluating the students' level of engagement without direct observation. Due to these measurement difficulties, in the method described below we make few claims about the student's level of engagement with the examples outside of class. We will distinguish between those problems that are suggested and those that are required, but it is certain that individual students will engage with the suggested problems in very different ways. As a result, we will make few judgments about the students' levels of engagement other than to distinguish cases where students were observed to actively engage with the examples from those where they were not.

Methods

3.1 Setting

In order to study the enacted curriculum, we collected data from an introductory abstract algebra class taught by a tenure-stream faculty member with an appointment in the department of mathematics and statistics at a mid-sized doctoral granting institution in the Northeast. The class was designed to be an introduction to the basic concepts of abstract algebra including groups, rings and fields. The class met three times per week for 70 minutes for 16 weeks. Students had frequent homework assignments, three in-class exams and one final exam.

The class had a mix of mathematics, mathematics education, and other STEM majors. Students ranged from sophomores to graduate students. The class had approximately 25 students. Although it was highly recommended that students had completed an introduction to proofs course, there were other ways to satisfy the pre-requisite (such as a lower-division linear algebra course) and as a result the student's prior proof-writing experience was quite varied.

3.2 Collecting data

In the lecture-based class we observed and video recorded 25 consecutive class meetings. We began with the class meeting immediately preceding the introduction of the formal definition of a mathematical group and ended with the definition of a factor group (quotient group). The camera was placed at the back of the class and pointed towards the board in order to best capture what the instructor said and what was written on the board. We also collected copies of all handouts given in class (homework assignments and exams) as well as tracked the assigned practice problems from the text.

All video data was digitized. Transana was used to code all incidents where an example or non-example was shown, constructed or analyzed in class. From the practice problems, homework assignments, and exams we recorded each instance where students were asked to work with a specific structure. Examples of problem-stems that indicated students were to consider a particular example were, "determine if the following form a group," or "Show that X is a group."

3.2 Analyzing data

We focused our analysis on the instructional uses of examples and non-examples of groups. In order to do this, we created an example log similar to Rasmussen and Stephan's (2008) argument log that included four columns to describe each example or non-example. The table was organized as follows:

- The first column listed each example or non-example of the particular construct; in this case the construct was an algebraic group.
- The second column listed the number of class meetings that had occurred since the formal definition of a group had been given that the example at hand demonstrated. A written homework assignment was coded as occurring on the day that it was assigned. This was meant to help describe its approximate position in the example neighborhood.
- The third column described the qualities of the example or non-example. In the case of examples, the third column described any additional qualities that the example possessed from a list that would be known to first semester algebra students by the midpoint of the semester (such as being a commutative group, a finite group, or a cyclic group). For non-examples, we described any properties of the construct that were missing as well as additional properties that the non-example possessed from the list above. This was meant to allow us to describe the range of permissible variation that was included in the enacted example space.
- The fourth column described the manner in which the example or non-example was made part of the classroom discourse as well as how the example was used. This was meant as a way to describe the example function.

Subsequently, we summarized the example space and the range of variability that was part of the enacted curriculum of the class as described in the section on our theory of the enacted example space. To describe the example neighborhood we present a narrative that describes each of the examples and when, in time, they were presented. We pay special attention to the first example and what other examples are presented soon after. We also described how the examples varied from one to the next, and, following Dienes (1963) suggestion, whether those examples varied along more than one dimension at a time. Finally, we give a description of the total variation in the qualities of the examples and non-examples that the students experienced. Collectively, we believe that these analyses allow us to describe the enacted example neighborhood.

After describing the example neighborhood, we analyzed all of the constructed examples. We first described the examples that were constructed by drawing on column four in the table. We also described the methods for constructing examples that the students was exposed to. The goal of describing the construction methods is to be able to predict the tools that students have to create their own examples, including novel ones.

In the third phase of the analysis, we drew on the fourth column of the table to describe the function of the example. The great majority of examples were used to exemplify a particular concept. There were relatively few examples that were used for any other function, but, for example, we did capture an instance of using an example to understand a theorem. We did this by describing the context of each example and how, if at all, the students or instructor made subsequent use of the example.

The fourth column also provided enough information to describe how the students engaged with the material, if at all, by describing how it became part of the classroom discussion and how it was used. From this column we described how students engaged with an example; for example, they might have all worked on the example during class. The example might have been assigned as part of a homework assignment to be submitted and, therefore, all students would have engaged with it in some form. Or the example might have been part of classroom discussion but without all students having been required to work on it. Finally, by drawing on the way that students engaged with examples, as well as all of the previous analyses, we drew inferences about the students' ability to learn and their likelihood of learning. In the case study presented below we will first present the collection of examples that the students experienced. Then, we will present subsections describing each of the aspects of the example space.

Data and Analysis

In the analysis that follows we present and analyze data immediately following the instructor's introduction of the formal definition of a group. These class periods drew on more examples per day than any other. They also featured a wider range of example and uses of example than others. We feel that the class periods directly following the introduction of the definition of a group are the best illustration of the instructor's teaching with examples.

The instructor introduced the definition of a by discussing the historical roots of algebra as a study of solving equations and asked the question, "In the case of equations of the form $a*x=b$, what properties do we require of the set and operation to be able to solve the equation?" He presented $x+3=5$ and $a*x=b$ and then drew out the required properties for the definition of a group.

4.1 Dr. P's Teaching

In the lecture-based course the instructor, Dr. P, defined a group and showed 6 structures within the same class period. Four of the structures were examples of groups. The first example was (Z, +), immediately followed by (Q, +). Both were asserted without proof based on previous work. He next proposed the rational numbers under multiplication as an example of a group, but immediately noted that 0 does not have an inverse. As a result, he stated that the structure is not a group. He then proposed a modification of the set to be the non-zero rational numbers and then verbally checked each of the required properties.

Next, he claimed that the non-zero real numbers under multiplication are a group for similar reasons. Dr. P. next proposed a set and arbitrary operation $(R-\{-1\},*)$ and asked, "How can we define * in such a way that this is a group?" He proceeded to define $a*b=a+b+(ab)$ and gave a verbal check that all properties held. Finally, he introduced the group $(Z_{12},+_{12})$ as an example. The entire sequence of discussion lasted approximately 25 minutes. The only other example introduced over the next two class periods was the more general case of the integers modulo *n*, $(Z_n,+_n)$. On the third day after introducing the definition of a group, Dr. P gave examples of the integers modulo 6 and 12 under appropriate additions and then spent 35 minutes proving that $(Z_n,+_n)$ is a group. He then introduced the concept of cancellation in groups via an example in the integers, $3+x=5$. He proceeded to solve the equation for *x* by substituting 3+2 for 5 and writing $x=2$. He asked, "Can we prove that we can always cancel by using our group axioms?"

On the fourth day after the definition of a group he asserted that 'mixed' cancellation ($a*b=c*a$ implies $b=c$) is not a logically valid claim and then stated, "we need some non-commutative group examples," but did not supply any. After asserting this, he asked, "when do groups have the same structure?" and drew on the examples of $(Z,+)$, $(Q,+)$, $(Z_{12},+_{12})$, and $(E,+)$ where *E* is the even integers. Dr. P gave a verbal explanation as to why $(E,+)$ is a group. He repeated his question, "when can these groups be the same?" He continued by saying, "We're in the driver's seat. We get to define this" (meaning, the sameness of groups). Dr. P then asked the class to vote on which of the examples they felt should be, "isomorphic" (he used the term without definition) or the same. The results were such that a plurality want $(Z,+)\cong(Q,+)$ and $(Z,+)\cong(E,+)$ while no students voted for any equivalence with the finite group or between $(Q,+)$ and $(E,+)$. In this case, the students showed some evidence of engaging with the examples.

On the fifth day after the definition, Dr. P drew on two different presentations of a group of 2 elements and showed that they were isomorphic to each other and to $(Z_2,+_2)$ both by rearrangement of the operations tables and renaming of the operation and then by using the formal definition of isomorphism of groups. Dr. P then showed that $(Z,+)\cong(E,+)$ and $(Z,+)$ is not isomorphic to $(Q,+)$. Finally, Dr. P assigned a set of practice problems that included a number of examples and non-examples.. Eight of the problems required students to determine if a set and operation formed a group or a subgroup of a specified group. Of those eight, five did not satisfy the properties of the group definition, with three of the non-examples being commutative and two non-commutative. The other three items included two commutative groups and one non-commutative group (the upper-triangular subgroup of the general linear group of *n* x *n* matrices with real number entries). The students were also to exhibit an abelian group of 1000 elements and to show that the complex numbers with a norm of 1 are not isomorphic to either the real numbers under addition or the non-zero real numbers under multiplication.

On the 7th day after the definition, Dr. P introduced the symmetric group on 3 elements and gave out the second required homework set. As part of that problem set, the students were to exhibit a group, *G*, which has a non-abelian subgroup, *H*, and then to justify their response. The students were highly likely to engage with this problem because it was homework to be submitted. Finally, on the 8th day, Dr. P spent the class proving the claim that every finite group of even order has an element of order two. He illustrated this claim with the examples of $(Z_{12},+_{12})$ and $(S_3,\circ)$.

Day 0	(Z, +), (Q, +),	
	$(Q,\cdot)$	Not a group
	$(Q^*,\cdot)$, $Q^* = \{q \in Q \mid q \neq 0\}$	Constructed
	$(R^*,\cdot)$	
	$(R-\{-1\},*)$	Constructed * such that the result is a group.
	$(Z_{12},+_{12})$	Introduced, not formally defined.
Day 3	$(Z_6,+_6)$, $(Z_{12},+_{12})$, $(Z_n,+_n)$	Formally defined, the first illustrated the general case
	$3+x=5$ implies $x=2$ by cancellation	Using an example from the integers to illustrate the idea of cancellation.
Day 4	(Z, +), (Q, +), $(Z_{12},+_{12})$, and $(E,+)$	When do groups have the same structure? When are they the same?
Day 5	3 different representations of $(Z_2,+_2)$	Showed they were isomorphic by re-arranging operation tables and via the formal definition
	Showed $(Z,+)\cong(E,+)$ $(Z,+)$ is not isomorphic to $(Q,+)$	Illustrating formal definition by showing that $(Q,+)$ is not cyclic.
	five did not satisfy the properties of the group definition, with three of the non-examples being commutative and two non-commutative. All infinite. The other three items included two commutative groups and one non-commutative group. All infinite exhibit an abelian group of 1000 elements show that the complex numbers with a norm of 1 are not isomorphic to the real numbers under addition or the non-zero real numbers under multiplication. Infinite and commutative	Practice problems
Day 7	$(S_3,\circ)$	Introduced and defined
	exhibit a group, G, which has a non-abelian subgroup, H,	As part of homework
Day 8	$(Z_{12},+_{12})$ and $(S_3,\circ)$.	Illustrating a claim that every finite group of even order has an element of order 2

4.1.1 Dr. P's enacted example neighborhood

The first example of a group that the instructor gives is the group of the integers under addition, an infinite cyclic group. The next three structures are all infinite and commutative, with the only variation the students experience being a change in either set or operation. The third proposed structure was a non-example, as the rational numbers under multiplication do not form a group, but was infinite and commutative. Thus, the students had early access to the idea that not all common sets and operations form a group. The fourth structure was a modification of the third, and, included the first new set, the non-zero rational numbers. The next two examples were similarly common sets, each missing one element, along with a commutative operation. The sixth operation was non-standard and offered students the opportunity to expand their example space by expanding their range of operations. The final example introduced on the same day as the definition was a finite, cyclic example, this was the first finite example that the students had experienced. By the third day after the introduction of the definition of a group Dr. P introduced a way to create finite cyclic examples of any size. On the fourth day after the introduction of the definition, Dr. P introduced a new example, the even integers under addition, as a way to illustrate the concept of isomorphic groups. This group is also an infinite and cyclic group with a known operation and reasonably common set. As a result, while the group is new to the students, it may not expand their example neighborhood in any significant way.

Thus, by the end of the fourth day, the students may have an example neighborhood that was populated by examples of groups of all possible sizes, that includes at least one non-standard operation, and included non-standard sets such as all of the real numbers except negative 1. Moreover, the students have access to the fact that not all sets and operations form groups. What the students have not yet experienced is a non-commutative example of a group, nor a non-commutative structure of any type. Some limitations to this enacted example space are evident: all the groups given were commutative, and only one non-example was proposed (and it was quickly revised into an example). It is possible that the students could develop significant misconceptions about the nature of groups, including the fact that all groups are commutative. Thus, we believe that this instructor did not offer his students a mathematically rich example neighborhood in the first four class periods. Furthermore, the non-example offered was insufficient to allow the students the opportunity to understand the limits of the concept of group. By this time Dr. P had started using the examples to illustrate concepts as well as propositions and had not given students an intellectual need to include non-commutative examples. We believe that this makes it likely that students would mistakenly believe that all groups are commutative. Similarly, because the students only saw one non-example and it was modified into an example, they may believe that all non-examples can be made into examples using the same modification strategy, or another one. Finally, we note that the students were expected to engage with the examples of $(Z, +)$, $(Q, +)$, $(Z_{12}, +_{12})$, and $(E, +)$ when they were asked to vote on which of them were isomorphic.

In the fifth class period after the introduction of the definition of a group the students experienced an additional variation in their example space. In that class period Dr. P assigned practice problems that included working with matrices, meaning that students were introduced to non-commutative structures. The first three non-commutative structures were also not groups, while only one of the non-commutative structures was a group. In order to complete all of the problems the students would also have needed to develop an understanding of the general linear group, another non-commutative group. All of the examples were infinite, and, just as Dienes (1963) suggested, they varied by one dimension at a time. On the seventh day after the introduction of the definition of a group Dr. P presented the permutation group on three elements. This was the first finite non-commutative group that the students experienced. Moreover, they were required to engage with such a group in the homework that Dr. P assigned that period.

The summary of the variation that students experienced is as follows: they saw commutative groups of all sizes, an infinite non-commutative group, and a 6 element non-commutative group. They saw cyclic groups of all sizes and infinite non-cyclic groups, including a proof that the rational numbers are not cyclic. The students experienced some artificial constraints on the variation of their example space; specifically relating to the sizes of non-commutative groups. Moreover, because the first three examples of non-commutative structures that students experienced, and, no non-commutative structures were introduced until five days after the definition of a group, we hypothesize that students are likely to hold the mistaken belief that all groups are commutative. In terms of how the variation was presented, with the exception of the introduction of the permutation group, it followed Dienes' (1963) recommendation about minimizing the amount of variation from one example to the next. Thus, we believe that the students had the opportunity to construct a well-defined example neighborhood, but, it was generally populated and defined by familiar structures.

4.1.2 Example construction

The instructor offered two different structures that were made into examples of the concept of group. The first was proposed, shown to not satisfy the axioms, and modified by removing the problematic element from the set. This gave the students the opportunity to learn one possible example construction strategy: remove problematic elements.

The instructor proposed another structure and asked what definition of an operation would make the structure a group. This called into being a non-standard operation. It gave the students access to a new range of operations as well as a new way to construct examples. They might have learned that they could begin with a set and then define an operation, standard or non-standard, that would make the structure a group. The instructor did not give the students a meaningful opportunity to learn how he created the operation, however, so it is not clear that the students would be able to immediately adopt the second example construction process. Given that the instructor offered two different methods, both illustrated with exemplars of the process, to construct new examples of groups, we believe that his instruction was of likely to help the students develop methods for constructing examples.

In their homework, the students were expected to exhibit a group with a non-commutative subgroup. Students may have attempted to construct another example using one of the two illustrated methods. But, they could have done this without constructing a new example by saying that a group is a subgroup of itself and then exhibiting $(S_3, \circ)$.

4.1.3 The function of examples

There were 4 different functions of examples illustrated in Dr. P's teaching. He used examples to illustrate definitions such as a mathematical group, isomorphic group, and an isomorphism of groups. He also used examples to instantiate statements of propositions after they had been introduced, such as when he drew on examples while discussing the claim that every finite group of even order has an element of order two. Dr. P used examples as a means to motivate the need for and introduce definitions of new constructs. For example, before he gave the definition of group or isomorphism he introduced the idea via a concrete example such as, "what properties of the set and operation are necessary in order to be able to solve this equation?" Finally, Dr. P used examples to motivate claims (he was not observed generalizing from examples, thus, we have chose the phrase motivate rather than create) before giving their formal statement such as when he drew on solving equations in the integers to motivate the question of whether cancellation is possible in groups.

With respect to the second aspect of the function of example, $(Z_{12},+_{12})$ was the most frequently cited example. It was used as an exemplar of two different concepts, a group, and when two groups are isomorphic. The group was also used to illustrate a claim about finite groups. Thus, we claim that in the enacted example space, $(Z_{12},+_{12})$, will possibly start to occupy a preeminent status. Given that it is the most frequently invoked, it would not be unreasonable for student to link example of $(Z_{12},+_{12})$ with the definition of a group as their concept image. This would be problematic due to the fact that $(Z_{12},+_{12})$ as a concept image would not allow students to reconstruct the definition of a group as it has additional properties.

Significance and Directions for Future Study

This paper makes four meaningful contributions to the research literature. The first significant aspect of this study is that it describes and analyzes the teaching of an undergraduate abstract algebra class. Before this, there were relatively few studies of undergraduate instruction in proof-based courses and none of an abstract algebra class. As this is a single case study, it is inappropriate to draw generalizations from it; yet, without a body of empirical evidence there is no basis for more theoretical work.

The second contribution of this study was to provide a finer-grained description of the example space. Mason and Watson (2008) described many of the criteria outlined in this paper, including knowing what can vary and what must stay constant, but we have added detail in the areas of example construction, such as being specific about what techniques for example construction students have access to. Similarly, we have added detail about function of examples, where the uses examples may be described. We believe that the construct of example function will give researchers a tool to analyze teaching as well as a way to better understand students' personal example neighborhoods. It will also provide more detail about students' concept image (Vinner, 1991) and their accessible example space.

The third contribution of this paper was a more detailed description of the example space as a tool for studying undergraduate teaching. Generally, "researchers' questions, methods, and analyses have not generally targeted what teachers say, do, and think about collegiate classrooms in an extensive or detailed way" (Speer et al., 2010, p. 105). These studies have failed to examine the teaching of any content in undergraduate classrooms in depth. In addition, there is an insufficient quantity of tools for doing so (Rasmussen and Marrongelle, 2006). This study created a new lens for analyzing undergraduate proof-based classes that gives insight into what students might learn. Although this technique was demonstrated in the setting of abstract algebra, we assert that it would be equally meaningful in any other proof-based undergraduate course.

This analysis provides a direction for future research, in addition to analyzing the enacted example functions. In particular, we propose to connect the analysis of the enacted example space with what students actually learn. This paper specifically described the opportunities students had in an algebra class to learn about examples of groups. We propose to investigate this question: does an increase in the number of examples discussed in class actually translate into students having richer individual example spaces?

Furthermore, while there is theoretical work describing the importance of students' example spaces and personal reflection from mathematicians that a rich example space helps in their work, there is no evidence that students at the advanced undergraduate level with a richer example space are more able to abstract, generalize or write proofs. We propose to investigate this correlation as well. We believe that generally a richer example space is better, but that the level of comfort and fluency with examples is also important. Neither of these concerns are adequately addressed in the analysis above.

Finally, we propose that this method of analysis gives undergraduate mathematics faculty and those who work with them a tool that may help reform instruction. More careful thought about the way that examples are presented, sequenced, and the functions that they are put to may cause faculty to think practice better pedagogy. Again, we present this as a possibility that would need further work in order to develop it.

References

Alcock, L. (2009). Mathematicians' perspectives on the teaching and learning of proof. *CBMS Issues in Mathematics Education, 16*, 73-100.

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

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

Courant, R. (1981). Reminiscences from Hilbert's Gottingen. *Mathematical Intelligencer*, *3*(4), 154–164.

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics, 33*, 283–299. doi:10.1023/A:1002999415887.

Dienes, Z. (1963). *An experimental study of mathematics-learning*. London: Hutchinson Educational.

Goldenberg, P. & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics, 69*, 183–194. DOI 10.1007/s10649-008-9143-3

Harel, G., & Fuller, E. (2009). Contributions toward perspectives on learning and teaching proof. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 355-370). New York, NY: Routledge.

Harel, G., & Sowder, L. (2007). Towards a comprehensive perspective on proof. In F. Lester (Ed.), *Second handbook of research on mathematical teaching and learning* (pp. 805-842). Washington, DC: NCTM.

Lakatos (1976). Proofs and Refutations. Cambridge: Cambridge University Press.

Larsen, S. and Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67,* 205-216.

Mason, J. & Watson, A. (2008). Mathematics as a Constructive Activity: exploiting dimensions of possible variation. In M. Carlson & C. Rasmussen (Eds.) *Making the Connection: Research and Practice in Undergraduate Mathematics.* (pp189-202) Washington: MAA.

Mejia-Ramos, J. P., and Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of the ICMI Study 19 conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 88-93), Taipei, Taiwan.

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

Rasmussen, C. & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics into instruction. *Journal for Research in Mathematics Education*, *37*, 388-420.

Rasmussen, C. & Stephan, M. (2008). A methodology for documenting collective activity. In A. E. Kelly & R. Lesh (Eds.). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching* (Pp 195-215). Mawah, NJ: Erlbaum.

Sfard, A. (1994). Reification as the birth of metaphor. *For the Learning of Mathematics, 14*(1), 44–55.

Speer, N., Smith, J., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99–114.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 25-41). Dordecht: Kluwer.

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

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. (2010, February). *The pedagogical practice of mathematicians: Proof presentation in advanced mathematics courses*. Conference on Research in Undergraduate Mathematics Education, Raleigh, NC. Retrieved from http://sigmaa.maa.org/rume/crume2010/Archive/Weber_Keith.doc.

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

Zodik, I. & Zaslavsky, O. (2008). Characteristics of teachers' choice of examples in and for the mathematics classroom. *Educational Studies in Mathematics, 69*, 165–182 DOI 10.1007/s10649-008-9140-6.

The Effects of Online Homework in a University Finite Mathematics Course

M. Axtell
University of St. Thomas

E. M. Curran
University of St. Thomas

Over the past 15 years, mathematics departments have begun to incorporate online homework systems in mathematics courses, touting benefits for students and instructors alike. However, the impact of web-based homework systems on student engagement, learning, and perception are poorly understood. This preliminary study seeks to add to this body of research by comparing the performance and experience of students taking an undergraduate finite mathematics course in a traditional paper/pencil homework format to that of students completing the same course, with the same instructor, using an online homework format. While statistically significant results comparing the two sections on learning outcomes were few, descriptive analysis yields consistent trends suggesting that student learning may be enhanced through online homework participation. However, despite potential positive impacts on learning, students in the online homework section were significantly less likely than their traditional homework counterparts to rate their course as "excellent" on end-of-semester course evaluation forms.

Key words: online homework, learning outcomes, effective practice

Introduction

It is an old, and often repeated, story that an emerging technology will serve as the catalyst of change within the classroom, and this is especially true in the math classroom. The appearance of cheap, hand-held calculators led to the disappearance of a host of computational techniques from high school and college math curriculli (as well as the slide ruler!), while the rise of computer algebra systems (CAS's) have, as can be seen in the Calculus Reform program of the 1980's and 90's, forced the teaching mathematician to re-examine the focus and goals of a mathematics course. More recently, web-based math homework systems have emerged and been widely adopted on many campuses across the nation at both the post-secondary and secondary level. The past 15 years have seen this nascent application blend the instant availability of the web with the computational power and flexibility of a CAS into a product that seeks to transform how students practice, and master, the computational techniques of a mathematics course.

There are many online mathematics systems that today's instructor may choose from. Some widely used examples are WebAssign, WeBWorK, SAGE and ALEKS, while there are a host of other systems that have been created in-house by, and for, specific institutions. Many of the commercial versions of this type of product have also begun to be customized to, and bundled with, specific textbooks. This linking of text to online homework product is perhaps a natural step in what may be an inevitable shift to the use of electronic texts and their support products in undergraduate mathematics courses.

The promise of these online systems is perhaps two-fold. For the student, these systems provide a setting to practice the computational techniques of a mathematics course and receive instant feedback concerning their results. Further, this opportunity to practice and receive feedback is available whenever the student desires. For the instructor, there are many perceived (and advertised) benefits. The instructor may allow their students to redo the problems as many, or as few, times as they wish. In some systems, the instructor may also choose the amount of feedback the student receives. And, perhaps the biggest draw is that the grading of the student work is done automatically. It is this benefit that might also appeal to the academic administrator – by requiring the student to purchase an online homework product the administrator need not provide funding for graders while freeing up some of their faculties' time for activities that are perceived to be more essential.

These seem very appealing. However, it is not clear if these online systems actually help student learning. This study grew from the desire of the two authors to investigate the pro's and con's of using online homework rather than a traditional paper/pencil homework setup.

A Review of the Literature

The process by which an individual comes to understand and master an idea or concept is what educational research is focused upon. As documented by von Glaserfield (2001) and Hauk & Powers (2006), understanding is constructed by the student in a self-regulated process. Self-regulation is a term used to describe the extent to which "individuals are metacognitively, motivationally, and behaviorally active participants in their own learning" (Zimmerman, 1994, p. 3). It is reasoned that students who are highly engaged with the content of a course will experience higher levels of achievement and performance. The use of homework, and in particular graded homework, in undergraduate mathematics education has been used to provide students with opportunities to engage in self-regulated learning. Multiple studies have shown that, in general, students that complete frequent homework assignments for instructor feedback and a grade demonstrate greater learning achievement in the mathematics classroom (Trautwein, Koller, Schmitz & Baumert, 2002; Cooper, Lindsay, Nye & Greathouse 1998; Paschal, Weinstein & Wahlberg 1984; and Wagstaff & Mahmoudi 1976). This result tends to be reliable even when prior knowledge and intelligence are controlled for within the context of the study (Trautwein, Koller, Schmitz & Baumert, 2002).

The results of the available research on the relationship between homework and learning outcomes, however, are far from straightforward or conclusive. Homework does not appear to have the same benefits for all groups of students. In fact, it has been found to be particularly beneficial for older students, students with higher socio-economic status, students with learning disabilities, and Asian-American students and less beneficial for students who are not members of those groups (Cooper, Robinson & Patall, 2006). Educational researchers have yet to determine how much homework is too much and how little is too little to effectively promote engagement and content mastery. Furthermore, in their 2003 paper, Trautwein and Koller pointed out systemic issues with many of the wide-scale homework studies of the past 25 years. Few have used randomization procedures or have adequate sample sizes; even fewer have differentiated between different aspects of the homework (e.g., type of homework, frequency of homework, length of homework) in their examinations of the impact of homework on learning outcomes. Much of what has been published on the topic may be called into question for flaws in study design that compromise the validity of the results (Trautwein, Koller, Schmitz & Baumert, 2002). These points are not mentioned to discredit earlier studies, but simply as a call for continued careful study of this deceivingly complex issue.

With these studies in mind, the use of an online homework tool to provide a frequent and monitored homework opportunity, with feedback, appears likely to contribute to positive learning outcomes for the student. In addition, a study by Hirsch and Weibel (2003) found that the use of an online homework system actually improved homework completion and homework success rates (as measured by scores based on the correctness of exercises). Thus, it could be reasoned, online homework systems would be expected to actually improve student learning outcomes when compared to a more traditional paper/pencil homework system.

It is not clear that this improvement actually takes place, however. In their 2006 study on the use of online homework systems in multiple sections of a college algebra course, Hauk, Powers, Safer and Segalla found no significant difference in learning outcomes between the 12 sections using an online tool and the seven sections using traditional paper and pencil homework. In a general Calculus course study, Hirsch and Weibel (2003) found a small, but statistically significant improvement in final exam performance by the sections that used an online homework tool over the traditional paper and pencil sections. However, in this study, all of the sections did paper and pencil homework and certain sections completed some of their homework online. Thus this study does not reflect the way many institutions use online homework systems. More recently, Zerr (2007) reports that the use of an online homework systems in Calculus I sections at one university led to improved student performance in learning outcomes as well as a higher level of student engagement and satisfaction with the course as compared to students in sections using a traditional paper and pencil homework setup.

As pointed out by Trautwein and Koller (2003), there is little research on the connection between different kinds of homework and student achievement in the context of a particular course or subject area. Moreover, the relationship between homework and course evaluation has yet to be thoroughly investigated by researchers. To add to the body of research on the impact of on-line homework in undergraduate mathematics education, the following questions were addressed in this study:

1. Does homework format (online vs. paper) impact learning outcomes?
2. How does homework format impact student perceptions of the class and instructor?

Study Design

This study was conducted at private, medium-sized, Liberal Arts institution in the upper Midwest. The institution supports many professional majors and serves approximately 6,000 undergraduate students, a majority of which would identify as Caucasian. In the spring of 2010, one of the co-investigators of this study taught two sections of Finite Mathematics – a freshman-level math class that satisfies the University's general education mathematics requirement. This course is taken almost exclusively by undergraduates whose majors do not require a Calculus course, and thus serves a somewhat weaker mathematical clientele.

For the study, one section used the online product, WebAssign, for its homework while the other section used a traditional paper and pencil format for homework. WebAssign was chosen since it is supported (in fact, a product of) the publisher of the textbook used in the course. This meant that both sections of students were doing the exact same problems throughout the study - with the minor caveat that the online students' problems would have small variations in the values used in a problem. (WebAssign randomizes certain values within most problems to reduce the chance that one student can simply copy another student's answers.) Both sections were given the same amount of time to complete each assignment and each assignment carried the same course value for both sections. Students in the online homework section were given three opportunities to correctly answer each homework question, while paper and pencil students were allowed to submit only one solution to each problem. However, students in the paper and pencil (henceforth, traditional) homework section were able to earn partial credit on their graded homework assignments. All assignments were graded (either by WebAssign or by a grader) and counted equally towards the course grade. Both sections took identical exams, and both sections met for 65-minutes on a MWF schedule. Thus, this study sought to create two sections identical in every controllable way with the exception of the homework system used.

The researchers were fortunate in that the two sections appeared almost identical in terms of incoming math ability (as measured by Math ACT score), though the two sections had significantly different ratios of male to female students (see Table 1). The Math ACT scores for the students ranged from 15 to 32 with a mean of 22.20 and a standard deviation of 3.25.

Table 1
Mean Math ACT and Gender Counts by Section

Section	Mean Math ACT	n_{males}	n_{females}
Traditional HW (paper/pencil)	22.18	23	6
Online HW	22.23	9	18

Results

The measure of learning outcomes chosen for this study is the student exam scores as well as individual question scores on the final exam. Recall that both sections took the same exams throughout the semester. Table 2 shows the two sections' performance on the three exams during the semester and the final, cumulative exam at the end of the semester. For the purposes of this study, a score of 70.0% or higher on an exam was considered a passing grade, while an exam score below 70.0% was classified as failing (though a course grade of 60.0% or higher ensured that the student passed the course).

From a descriptive point of view (see Table 2 and Figure 1), both sections appear to perform equivalently on Exam I, with similar proportions of students "passing" the exam. However, the on-line section appears to out-perform the traditional paper/pencil homework section on Exams II, III and IV with a greater proportion of students passing.

Table 2
Proportion of Students Passing/Failing Exams by Section

	Traditional[a]		Online[b]	
Outcome	*Pass %*	*Fail %*	*Pass %*	*Fail %*
Exam I	89.7	10.3	88.9	11.1
Exam II	75.9	24.1	88.9	11.1
Exam III	75.9	24.1	92.6	7.4
Final Exam	69.0	31.0	77.8	22.2

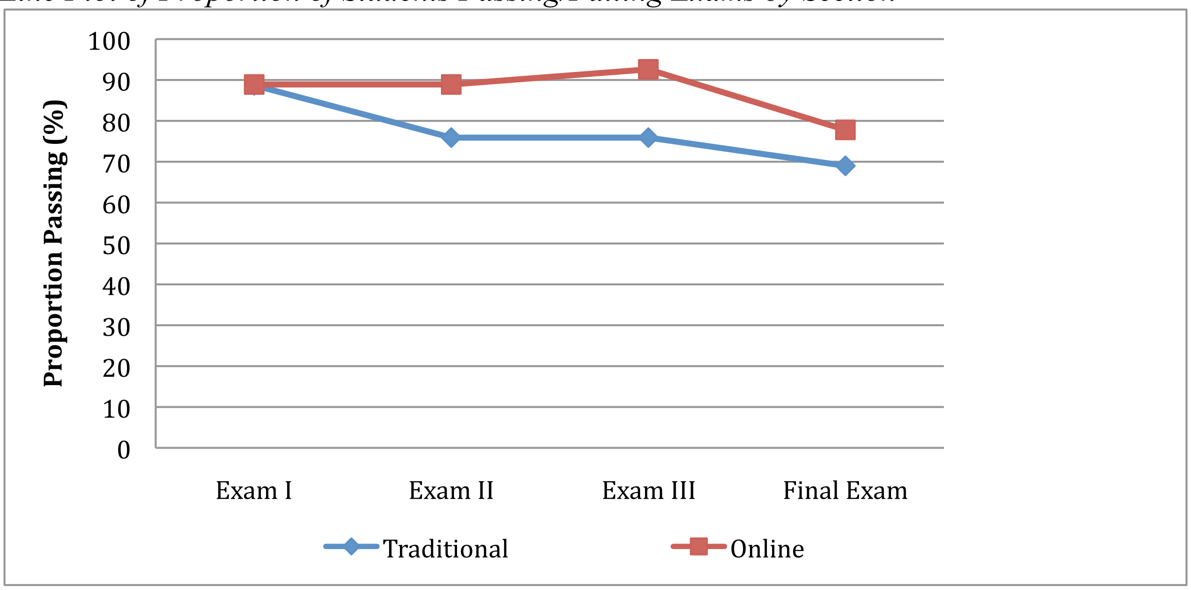
[a]$n = 29$ [b]$n = 27$

Inferential z-procedures for comparing the proportion of students passing in the traditional section to the proportion of students passing in the online section were used to identify significant differences, if any, between the sections. Despite apparent descriptive differences between the two sections on Exams II, II and IV, no statistically significant differences were found between the two sections in terms of the proportion of students passing each exam.

The second comparison looked at the difference between the traditional section and the online section in terms of passing/failing individual questions on the cumulative final exam. As in the previous analysis, a score of 70.0% or higher on an individual question was considered a passing mark, while a score below 70.0% was deemed failure for purposes of this study. The passing percentages on the final exam's 16 questions are given in Table 3 for each section.

Figure 1

Line Plot of Proportion of Students Passing/Failing Exams by Section



Descriptive analysis reveals that both sections appear to perform equivalently on Questions 5 – 6, 9 and 15 with highly similar proportions of students "passing" these items (see Table 3 and Figure 2). However, the online section appears to perform better than the traditional homework section on ten of the individual items (Questions 1 – 2 , 4, 8, 10 – 14 and 16) on the final exam; the traditional section performs better, as a whole, on just two items (Questions 3 and 7).

Table 3

Proportion of Students Passing/Failing Individual Questions on Final Exam

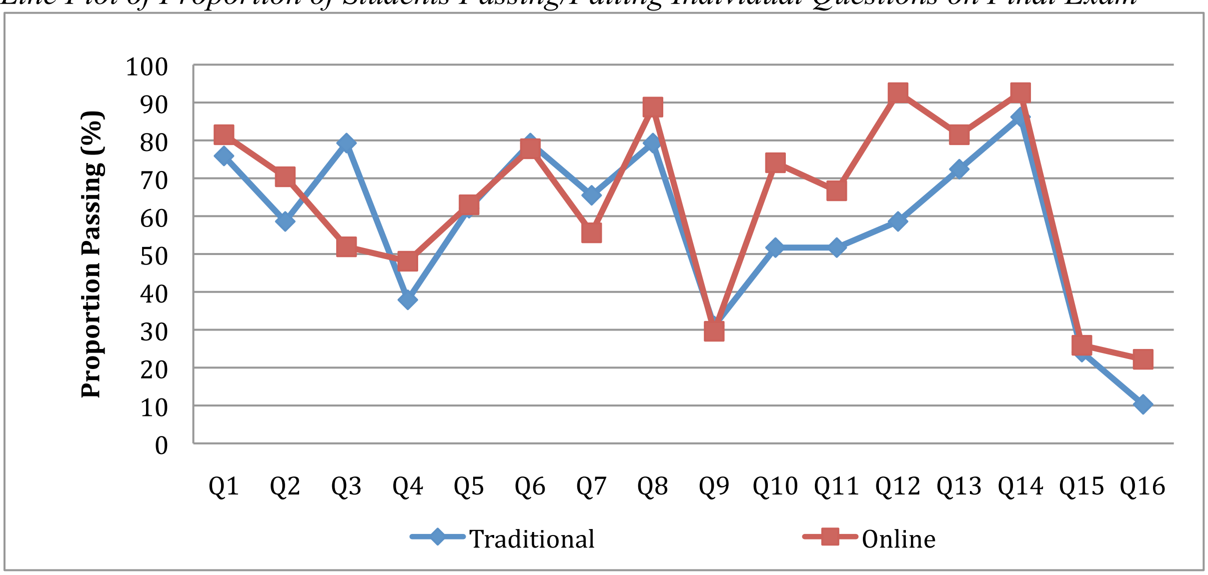
Outcome	Traditional[a]		Online[b]	
	Pass %	*Fail %*	*Pass %*	*Fail %*
Question 1	75.9	24.1	81.5	18.5
Question 2	58.6	41.4	70.4	29.6
Question 3	79.3	20.7	51.9	48.1
Question 4	37.9	62.1	48.1	51.9
Question 5	62.1	37.9	63.0	37.0
Question 6	79.3	20.7	77.8	22.2
Question 7	65.5	34.5	55.6	44.4
Question 8	79.3	20.7	88.8	11.2
Question 9	31.0	69.0	29.6	70.4
Question 10	51.7	48.3	74.1	25.9
Question 11	51.7	48.3	66.7	33.3
Question 12	58.6	41.4	92.6	7.4

Question 13	72.4	27.6	81.5	18.5
Question 14	86.2	13.8	92.6	7.4
Question 15	24.1	75.9	25.9	74.1
Question 16	10.3	89.7	22.2	77.8

$^{a}n = 29$ $^{b}n = 27$

Again, z-procedures for comparing the proportion of students passing individual questions on the final exam in the traditional section to the online section were used to identify the existence of significant differences between the sections. This analysis revealed three interesting outcomes for consideration. First, there are only two questions where one section outperformed the other in a statistically significant manner, Question 3 ($z = 2.25$, $p = .025$) and Question 12 ($z = -3.25$, $p = .001$), each of which will be discussed in detail below. Second, the traditional section outperformed the online section on Question 3, while the reverse was true on Question 12. Finally, though only Questions 3 and 12 exhibited statistically significant differences, the data indicate a descriptive trend in which the online section appears to be outperforming the traditional section on most of the questions – mirroring a trend seen in the data concerning exam pass/fail rates – though again, not to a statistically significant degree (See Figure 2).

Figure 2
Line Plot of Proportion of Students Passing/Failing Individual Questions on Final Exam



There were just two final exam questions in which statistically significant differences occurred, Questions 3 and 12. Question 3 saw the traditional section, as a whole, outperform the online section. This question can be viewed as a computationally 'expensive' test item involving the process of matrix multiplication:

Final Exam Question 3: Determine whether ▯=−13−134301−123−13−23 is the inverse of the matrix ▯=32322121−2.

This question requires careful and repeated calculations to work accurately. It is conjectured that the online section was not used to doing such work on paper. Further, this question appeared in a homework assignment and the online section had to answer either Yes or No, and they had three chances to get the right answer.

The other question on the final exam where statistically significant differences occurred between the two sections was Question 12:

Final Exam Question 12: Mr. Bean has three grocery stores in his neighborhood. There is a 25% chance he will go to store A, a 40% chance he'll go to store B, and the rest of the time he goes to store C. At store A there is a 40% chance he'll buy pears. At store B there is a 50% chance he'll buy pears, while at store C there is a 20% chance he'll buy pears.

a) What is the probability that Mr. Bean buys pears on a trip to the grocery store?
b) What is the probability that Mr. Bean went to grocery store B given that he bought some pears?
c) What is the probability that Mr. Bean buys pears if he goes to store C?

On this question, the online section outperformed the traditional section in a statistically significant way. This question requires the students to know how to compute an *a priori* conditional probability and an *a posteriori* conditional probability *and* to know which is called for in various situations. Perhaps because students in the online section received immediate feedback when they got this type of item wrong during homework completion (followed by two opportunities to correct their answers) they were more prepared than students in the traditional homework section for this type of item on the exam. The immediate feedback provided by WebAssign may have served to better highlight the differences between these two related probabilities.

This study also sought to determine any differences in perceptions of the instructor and course exhibited by the two sections. Student responses on the course evaluation form given at the end of the semester were used for this. The university in this study uses the IDEA forms, which are widely used throughout the nation. Both sections used the Long Form of this tool to evaluate the instructor and the course. Some of the relevant questions examined for the purpose of the study are listed below:

IDEA Items – Instructor:

1. Displayed a personal interest in students and their learning.
2. Found ways to help students answer their own questions.
3. Scheduled course work in ways which encouraged students to stay up to date.
4. Explained reasons for criticisms of student academic performance.
5. Stimulated students to intellectual effort beyond that required by most courses.
6. Gave tests, projects, etc. that covered the most important points of the course.
7. Provided timely and frequent feedback on student work.
8. Asked students to help each other understand ideas or concepts.
9. Encouraged student-faculty interaction outside of class.

IDEA Items – Course:

1. Amount of work in non-reading assignments as compared to other courses I've taken.
2. Difficulty of subject matter.
3. I worked harder on this course than most I've taken.
4. As a result of taking this course, I have more positive feelings toward the field of study.

IDEA Items – Overall:

1. Overall, I rate this teacher as excellent.
2. Overall, I rate this course as excellent.

Both sections appear to have felt quite positive about the instructor with the vast majority of students in both sections rating the instructor as excellent, and no statistically significant difference identified for that particular IDEA item. A statistically significant difference was identified, however, between the two sections' ratings of the course as excellent. Nearly 77.0% of the traditional homework section rated the course as excellent, while only 45.0% of the online section rated the course as excellent. In other words, despite a trend for somewhat higher achievement in the online section, the online section was significantly *less* likely than the traditional section to rate the course as excellent ($z = 2.30$, $p = .021$).

Summary

Overall, inferential analyses of achievement differences between the traditional and online homework sections of an undergraduate finite mathematics course indicate that the two methods for promoting student engagement and enhancing course outcomes are relatively equivalent. No significant differences were found between the sections on exam performance and only two significant differences, albeit with opposite results, were identified on individual final exam question performance. The results of this study may yield support for traditional approaches to homework for helping students master computationally intensive processes and problems in undergraduate mathematics courses; however, the results of this study may also provide support for the benefits of immediate feedback provided through web-based homework systems in enhancing engagement and learning. As is consistent with the current research on the impact of homework (in all of its forms) on learning, these results reinforce the notion that homework is not a one size fits all solution to the issue of student engagement and content mastery.

The inferential analyses completed in this study may actually only tell a portion of the story. While few statistically significant results were identified, there was a general trend for the online homework students to perform better, as a class, than students participating in the traditional homework section. These results may lend support to the assertion that because online homework systems allow students multiple opportunities to identify a solution to a problem, they may promote greater levels of student engagement and thus greater student achievement. This effect in the population may be small to moderate and may require a larger sample, yielding greater statistical power, in order to identify it reliably. The general trend of online homework systems fostering small to moderate advantages over traditional homework approaches is consistent with much of the recent literature on the topic.

It is important to note that not all of the results of the initial analyses are consistent with what has been published in the literature. The results of this study suggest that students using an online homework system may actually be significantly *less* likely to hold positive impressions of the course (despite a tendency to achieve better learning outcomes as a class) than their traditional homework peers. Anecdotal evidence suggests that students may initially experience frustration in learning how to use the online homework system and this may account for some of the negative perceptions of the course. Determining whether this result is reliable, and if so the reasons for this result, may be important for mathematics educators whose jobs and salaries depend, at least in part, on course evaluations completed by students.

This preliminary report marks the completion of the first part of this study. Due to the relatively small number of students currently in the study, it is important to interpret these findings as preliminary and with caution. It is hoped that a continuation of this study will enlarge the sample size sufficiently to determine whether some of these perceived trends and differences are, in fact, reliable trends and differences. In the fall of 2011, the researchers will again study multiple sections operating under the same conditions as the two described in this paper. Extending this study in a longitudinal nature and increasing the sample size will allow the researchers to further examine the impacts of online vs. traditional homework, the kinds of students that may benefit (or be negatively impacted) from regular online homework assignments, and the impact of web-based homework systems on student perceptions of a course and instructor.

References

Cooper, H., Lindsay, J.J., Nye, B. & Greathouse, S. (1998). Relationships among attitudes about homework, amount of homework assigned and completed, and student achievement. *Journal of Educational Psychology, 90 (1)*, 70-83.

Cooper, H., Robinson, J.C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research. *Review of Educational Research, 76*, 1-62

Hauk, S., Powers, R., Safer, A., & Segalla, A. (2006). A comparison of web-based and paper and pencil homework on student performance in college algebra. Retrieved November 14, 2010, from http://hopper.unco.edu/hauk/segalla/WBWquan_060307.pdf

Hirsch, L., & Weibel, C. (2003). Statistical evidence that web-based homework helps. *Focus, 2*, pg. 14.

Paschal, R., Weinstein, T., & Walberg, H. (1984). The effects of homework on learning: A quantitative synthesis. *J. of Educational Research, 78 (2)*, 97-104.

Trautwein, U., & Koller, O. (2003). The relationship between homework and achievement – still much of a mystery. *Educational Psychology Review 15(2)*, 115-145.

Trautwein, U., Koller, O., Schmitz, B., & Baumert, J. (2002Do homework assignments enhance achievevment? A multilevel analysis in 7th-grade mathematics. *Contemporary Educational Psychology 27, 26-50.*

Von Glasersfeld, E. (2001). Radical constructivism and teaching. *Revue Canadienne de l'enseignement des sciences, des mathematiques et des technologies, 1 (3)*, 211-222.

Wagstaff, R., & Mahmoudi, H. (1976). Relations of study behaviors and employment to academic performance. *Psychological Reports, 38*, 380-382.

Zerr, R. (2007). A quantitative and qualitative analysis of the effectiveness of online homework in first-semester Calculus. *Journal of Computers in Mathematics and Science Teaching, 26 (1)*, 55-73.

Zimmerman, B. (1994). Dimensions of academic self-regulation: A conceptual framework for education. In D. H. Schunk & B. Zimmerman (Ed.s) *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

A REPORT ON THE EFFECTIVENESS OF BLENDED INSTRUCTION IN GENERAL EDUCATION MATHEMATICS COURSES

Anna E. Bargagliotti
The University of Memphis
brggltti@memphis.edu

Fernanda Botelho
The University of Memphis
mbotelho@memphis.edu

Jim Gleason
University of Alabama
jgleason@as.ua.edu

John Haddock
The University of Memphis
jhaddock@memphis.edu

Alistair Windsor
The University of Memphis
awindsor@memphis.edu

Abstract

Despite best efforts, hundreds of thousands of students are not succeeding in postsecondary general education mathematics courses each year. Using data from 11,970 enrollments in College Algebra, Foundations of Mathematics, and Elementary Calculus from fall 2007 to spring 2010 at the University of Memphis, we compare the impact of the Memphis Mathematics Method (MMM), a blended learning instructional model, to the traditional lecture teaching method on student performance and retention. Our results show the MMM was positive and significant for raising success rates particularly in Elementary Calculus. The results also show the MMM as a potential vehicle for closing the achievement gap between Black and White students.

Key Words: Calculus, algebra, general education mathematics, retention, student performance

In the U.S., students who pursue a postsecondary baccalaureate degree are required to complete at least one general education mathematical science course. Low student success rates in these courses are pervasive, and efforts to improve student learning and success rates are crucial. National recognition of the poor success rates in such courses has resulted in a series of proposed reform models over the past two decades, usually as curricular reform or delivery reform. Numerous reforms have focused on technology. They have included attempts to change instructional delivery methods by training students to use technology to solve problems (Lavicza, 2009; Heid & Edwards, 2001; Smith, 2007), using technology as an instructional tool (Peschke, 2009; Judson & Sawada, 2002; Caldwell, 2007; Fies & Marshall, 2006), and using a technology based assessment system (Zerr, 2007; Nguyen, Hsieh, & Allen, 2006; Vanlehn, et al., 2005).

In this paper, we report results comparing the impact of the Memphis Mathematics Method (MMM), a blended learning instructional model, to the traditional lecture teaching method on student performance and retention in general education mathematics courses at the University of Memphis (UM). The comparison includes a total of 11,970 enrollments in College Algebra, Foundations of Mathematics, and Elementary Calculus from fall 2007 to spring 2010. Results indicate that the MMM is effective in increasing student achievement and retention.

Background

There is a general belief that instructional delivery methods directly affect students' learning environment and hence indirectly affect student performance. For example, an environment in which students actively participate and engage in learning likely creates rich opportunities for deep learning of mathematics (Schoenfeld, 1994; Henningsen & Stein, 1997). Moreover, there is mounting evidence that integrating technology in undergraduate instruction positively associates with student achievement (Alldredge & Brown, 2006; O'Callaghan, 1998) and attitudes (Hauk & Segalla, 2005; Cretchley, Harman, Ellerton, & Fogarty, 2000). Similarly, research confirms that computer instruction may be as or more effective than traditional classroom instruction due to the self-paced and individualized nature of the instruction (Means, Olson, & Singh, 1995; Barrow, Markman, & Rouse, 2009; Liao, 2007).

The MMM is designed to reflect the current understanding of the effective use of technology in the classroom both to create an active blended learning environment that is aligned with cognitive principles and to allow for more effective management of the classroom and instructor time. In addition, utilizing the features of MyMathLab software, the MMM aims to more effectively engage students with mathematics in a non-threatening manner that bolsters student success and confidence.

Framework

The following diagram represents the conceptual framework driving the development of the MMM.

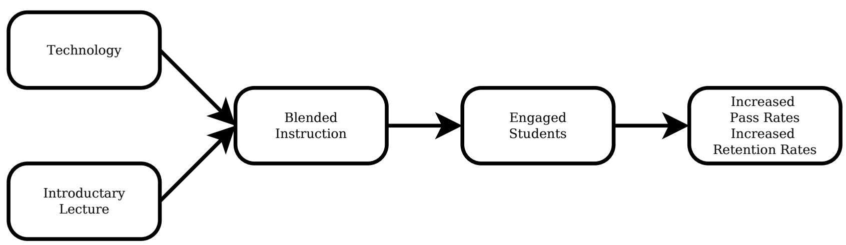


Figure 1. Framework used to guide the MMM model

In general, technology is believed to have a positive impact on student learning in mathematics. Many studies conducted in K–12 environments have reported significant gains in learning or learning speed (Koedinger et al., 1997; Fletcher, 2003; Anderson et al., 1995) when technology is incorporated into instruction. At the postsecondary level, studies have shown an increase in student success and learning when technology is employed in the classroom (O'Callaghan, 1998; Yaron, Cuadros & Karabinos, 2005; VanLehn et al., 2005; Ringenberg & VanLehn, 2006; Matsuda & VanLehn, 2005).

The implementation of technology through blended instructional strategy aligns with a variety of theoretical orientations that appeal to cognitive flexibility (Spiro, Feltovich, Jacobson, & Coulson, 1992), integrating abstract and concrete representations of concepts (Pashler, et al., 2007), embodied cognition (De Vega, Glenberg, & Graesser, 2008), combining inquiry and knowledge building (Mayer R. E., 2003), and other perspectives in the constructivist tradition. Recently, researchers have begun to make recommendations as to the appropriate proportion of student-centered and teacher-guided instruction (Chi, Siler, Jeong & Hausmann, 2001). For example, Mayer (2004) suggests that a blend of instructional methods be used rather than pure student-centered discovery. Using technology in the classroom can create a student-centered, active learning environment (White & Frederiksen, 1998; National Research Council, 2000; Fletcher, 2003). Computers and tutoring software are particularly effective tools in increasing learning (Sandholtz, Ringstaff, & Dwyer, 1997; Lowther, Ross, & Morrison, 2003; Smaldino).

This evidence suggests that a blended instructional method—technology coupled with guided lecture—may be ideal for increasing learning and success.

The MMM utilizes MyMathLab software to deliver the technology component of the general education math courses. MyMathLab can provide students with instant feedback for their work which research has shown leads to improved student achievement (Brooks, 1997; de La Beaujardiere et al., 1997; Khan, 1997). In addition, MyMathLab offers student aid features that align with elements identified in the literature as fostering increased student learning and understanding. These five learning aids are: (1) step-by-step worked solution of a similar problem, (2) video example, (3) just-in-time, (4) view an example, and (5) ask my instructor. First, the "step-by-step worked solution of a similar problem" tool can help students scaffold the content being covered in the problem which can help promote a deep understanding of content in computer-based training (VanLehn, 2006). Second, the multimedia tool "video example" capitalizes on the advantages of multiple media and modalities in improving learning and memory (Mayer, 2005; Pashler et al., 2007). Third, the availability of the electronic textbook while working through a problem allows a learner to access information "just-in-time" for achieving learner goals during problem solving (Rouet, 2006). MyMathLab directs students to the appropriate location in the textbook for the topic they are working through. This retrospective learning strategy allows students to read text when it is needed, which has been shown to increase learning of difficult content (Bransford & Schwartz, 1999). Fourth, "view an example" guides the student through example problems with solutions, a technique that is compatible to the research of Sweller on worked-out examples (Sweller & Chandler, 1994). Fifth, the "ask my instructor" conversational aid is comparable to intelligent tutoring systems that help students learn by holding a conversation in natural language (VanLehn et al., 2007). Collectively, these tools define MyMathLab as interactive content delivery software that aligns with cognitive principles of learning and curriculum in a blended instructional setting.

The Memphis Mathematics Method

The MMM substitutes traditional lecture-style instruction with a brief introduction of a topic followed by a laboratory session requiring students to complete classroom-based assignments using MyMathLab software. The MyMathLab software was selected because it offers student aid features that align with elements identified in the literature as fostering increased student learning and understanding.

Instructors employing the MMM begin each class with a 25-minute lecture followed by a problem-solving session using MyMathLab. During the short lecture, instructors introduce basic concepts and provide examples that emphasize the use of mathematical techniques to solve problems motivated by other sciences. Each lecture contains a list of objectives, a few illustrative examples, and mathematical problems for discussion during the presentation. Over the course of a 15-week semester, students log 30 hours of class time practicing problems on MyMathLab. In addition to its use as an instructional tool, instructors use the MyMathLab learning environment for course management and grading.

The remaining class time is dedicated to solving problems using the MyMathLab software. The problems chosen are a combination of review questions from the previous class period and problems directly related to the concepts presented in the introductory lecture. The instructor and an assistant, typically an advanced undergraduate student or a graduate student, are available during the class period to provide individual help and answer technical questions.

Final grades are computed as a weighted sum of all the points earned throughout the semester, including attendance, in-class lab assignments, tests, quizzes, and a final exam. Students complete proctored tests and the final exam online in the instructional lab.

Data and Methods

The MMM intervention was piloted at UM in 2007 in a specialized Developmental Studies Program in Mathematics (DSPM) College Algebra course, which combined a remedial Intermediate Algebra course with a regular College Algebra course. Students were eligible for the DSPM course only if their ACT scores would have required them to take remedial Intermediate Algebra.[i] In 2008, based on positive student outcomes during the initial pilot, UM expanded the use of the MMM to regular sections of College Algebra; regular and DSPM sections of Foundations of Mathematics; and regular sections of Elementary Calculus.

Regular courses of Elementary Calculus; both DSPM and regular courses of College Algebra; and DSPM and regular courses of Foundations of Mathematics have used the MMM. This study includes data from fall and spring semesters from fall 2007 to spring 2010. There were 11,970 enrollments in the sections across the three courses. Of these, 10,424 enrollments were in regular sections while 1,546 enrollments were in DSPM sections.

College Algebra at UM covers basic algebraic tools and concepts with an emphasis on developing computational skills necessary for success in subsequent mathematics courses. During the course of the study, there were 4,777 enrollments in this course. Of these, 3,668 were taught in a conventional setting, of which 157 enrollments were in DSPM sections, and 3,511 were in regular sections. A total of 1,010 enrollments were in DSPM sections taught using MMM and 99 enrollments were in regular sections taught using MMM.

Foundations of Mathematics provides instruction in basic logic and problem-solving skills. Students who enroll in this course are typically non-STEM majors who choose this course to fulfill their general education requirement. From fall 2007 to spring 2010, there were 3,986 enrollments in this course. Of these 3,525 were taught traditionally, 264 were taught traditionally in DSPM courses, 461 were taught using MMM, and 115 were taught by MMM in DSPM courses.

Elementary Calculus introduces the tools of differential calculus with emphasis on solving problems motivated by the social and life sciences, economics, and business. From fall 2007 to spring 2010, there were a total of 3,207 enrollments in this course. Throughout the duration of this study, 2,729 enrollments were taught traditionally, and 478 were taught using MMM. Since completing College Algebra or having a sufficiently high ACT or SAT score were prerequisites for Elementary Calculus, there were no remedial sections of Elementary Calculus offered

Dependent variables.

To gauge student success in the three courses, we define an indicator variable "success" coded as 1 if a student obtains a grade of C or above and 0 if they obtained a grade of D or F, or withdrew from the course. The variable success thus combines the effects of changes in pass rate and changes in dropout rate.

In addition, we are interested in separately determining the effects of the MMM pedagogy on dropout rates. We define an indicator variable "dropout" coded as 1 if a student dropped out of a course and 0 if a student completed the course. Success and dropout serve as our dependent variables in this study.

Independent variables.

We include the student's gender, the student's racial/ethnic background (White, Black, Hispanic, and Other), and the student's prior mathematics knowledge as measured by their ACT math score, as three independent variables in the analysis. In addition, we control for whether a student is repeating the course and define an indicator variable "redo" coded as 1 if a student has attempted the course before and 0 if this is their first attempt. Also, an indicator variable for whether a student was exposed to the conventional or to the MMM pedagogy is included in the analysis. Table 1 provides the descriptive statistics.

Table 1. *Descriptive Statistics of the Variables*

Variable	N	Mean	S. D.	Min	Max
Independent Variables					
ACT Math Score	9984	19.44	3.82	9	35
Redo	11970	0.15	0.35	0	1
Female	11970	0.59	0.49	0	1
Race					
White	6,059				
Black	5,354				
Hispanic	210				
Other	311				
Teaching Method					
Traditional	9,501				
MMM	923				
DSPM - Traditional	421				
DSPM - MMM	1,125				
Dependent Variables					
Dropout	11970	0.13	0.34	0	1
Succeed	11970	0.54	0.50	0	1

Estimation approach.

To estimate the effects of the MMM on student success and dropout rates in these courses, we fit a total of 6 regressions – three interactive models for remedial courses, and three interactive models for non-remedial courses.

To model the success rate and the dropout rate for both DSPM and regular courses, we fit logistic regressions for each of the three courses separately. Thus, we estimate the following:

$$logit(p_i) = ln(p_i/1\text{-}p_i) = a + X_{1i}\beta_1 + X_{2i}\beta_2 + X_{3i} * X_{2i}\beta_3 + u_i$$

where p_i is either the probability of student i succeeding or dropping out, X_1 is a vector of observed student characteristics (gender, racial/ethnic background, ACT score, and redo), β_1 is the associated coefficient vector, X_2 is a dummy variable for whether student i was exposed to the MMM pedagogy, and β_2 is its associated coefficient, X_3 is the vector of dummy variable for the different racial/ethnic backgrounds of students and β_3 is the associated coefficient vector for the interactive term.

Results

Descriptive results. Table 1 illustrates that of the 11,970 enrollments, 5,530 ended in a passing grade reflecting a 54% success rate over the three courses. Of the 11,970 enrollments, 1,596 ended when the student withdrew from the course.

To begin exploring whether the MMM is effective in increasing student success and retention in core general education mathematics courses, we first examine descriptive breakdowns of success rates and dropout rates by teaching pedagogy. Table 2 provides a numeric breakdown of student success and dropout over the study period. Overall, the tables illustrate that students in the MMM classrooms withdraw less and perform better.

Table 2. *Cross tabulates of succeed and dropout by teaching method for each course*

	Foundations of Mathematics				College Algebra			
	Traditional	MMM	DSPM Traditional	DSPM MMM	Traditional	MMM	DSPM Traditional	DSPM MMM
Not succeed	1,713	49	68	414	1,540	185	119	50
Succeed	1,909	57	89	647	1,823	181	146	66
Dropout	459	12	21	104	429	33	30	10
Not dropout	3,163	94	136	957	2,934	333	235	106

For every course, the percentage of students who withdrew from the MMM classes is lower than in the traditional classes. For example, 17.7% of students in traditional Elementary Calculus withdrew while only 8.2% withdrew from the equivalent MMM courses. In College Algebra, students in MMM classes drop out at a rate of approximately 9%. Their equivalent conventional teaching method courses have dropout rates of 12.8% for regular students and 11.4% for DSPM students.

With respect to performance, more students are succeeding in MMM classes than in traditional classes. In DSPM courses for Foundations of Mathematics, for example, 56.7% of students received passing grades, while 60.7% passed the equivalent MMM classes. Furthermore, a striking difference of grades across instructional methods is seen in Elementary Calculus. Approximately 49% of students in traditional courses passed while about 72% passed when exposed to the MMM teaching methodology.

Additionally, in Table 3, we compare the numeric breakdowns of student performance and retention by racial/ethnic background for each course, and find that racial disparities between Black and White students in performance seem to be greatly reduced in the MMM classes.

Table 3. Cross tabulates for race by success & dropout by teaching method

	College Algebra	Foundations of Mathematics	Elementary Calculus

	Black	Hispanic	Other	White	Black	Hispanic	Other	White	Black	Hispanic	Other	White
Not Succeed												
Traditional	855	36	49	730	817	30	19	622	676	18	29	665
MMM	24	2	0	18	109	5	0	59	43	2	2	86
DSPM-Traditional	55	1	1	11	85	2		32				
DSPM-MMM	295	2	3	97	30	3	0	17				
Succeed												
Traditional	603	38	59	1141	604	27	39	1103	356	20	71	894
MMM	21	0	3	31	71	6	3	93	134	6	14	191
DSPM-Traditional	44	4	1	40	84	3		58				
DSPM-MMM	412	5	17	179	36	0	1	28				
Dropout												
Traditional	194	9	13	228	201	10	5	203	231	8	9	241
MMM	6	0	0	5	20	2	0	9	12	1	2	25
DSPM-Traditional	17	1	0	3	21	0		9				
DSPM-MMM	74	1	0	26	4	0	0	6				
Not Dropout												
Traditional	1264	65	95	1643	1220	47	53	1522	801	30	91	1318
MMM	39	2	3	44	160	9	3	143	165	7	14	252
DSPM-Traditional	82	4	2	3	148	5		81				
DSPM-MMM	633	6	20	26	62	3	1	39				

Across all three regular courses, Black students pass at a rate of 39.9% when taught using traditional pedagogy compared to 56.2% when using the MMM. This difference is staggering. In DSPM courses, Black students dropout at a rate of 10% for the MMM method compared to a rate of 14% for traditional teaching.

Looking at these figures within each course, Table 3 reveals that this pattern of improvement persists. For example, in traditional DSPM College Algebra, 44% of Black students received passing grades compared to 78% of White students; that is, there is a 43% differential between Black and White students. In the equivalent MMM courses, however, this differential is reduced to 6%. In Elementary Calculus, the racial disparity between Blacks and Whites is completely erased with the MMM pedagogy with 75.7% of Black students and 68.9% of White students receiving passing grades. These results identify the MMM as a potential vehicle for decreasing the achievement gap.

In addition to reducing racial disparities in the passing rate, racial disparities in the withdrawal rate are also decreased. In traditional Elementary Calculus, 22.4% of Black students dropped compared to 15.4% of White students, while in the MMM calculus courses, only 6.8% of Blacks withdrew compared to 9% of Whites. These relationships are further examined in the following section using regression.

Regression results. The regression output for success is presented in Table 4 and the output for retention is presented in Table 5.

Table 4. *Logistic Regression of Success Against Explanatory Variables*

	Regular			DSPM	
Variables	Foundations	Algebra	Calculus	Foundations	Algebra
Female	1.18*	1.38***	1.47***	1.89**	1.33**
	(0.102)	(0.108)	(0.128)	(0.546)	(0.185)
ACT Math Score	1.13***	1.17***	1.11***	1.19**	1.15***
	(0.014)	(0.015)	(0.013)	(0.106)	(0.062)
Redo	0.66***	0.36***	0.92	1.06	0.68*
	(0.080)	(0.042)	(0.104)	(0.486)	(0.152)
Black	0.62***	0.71***	0.51***	0.84	0.30***
	(0.056)	(0.061)	(0.054)	(0.269)	(0.132)
Hispanic	0.64	0.66	0.92	0.54	0.91
	(0.184)	(0.184)	(0.359)	(0.518)	(1.213)
Other	0.98	0.81	1.79**		0.24
	(0.276)	(0.169)	(0.422)		(0.325)
MMM	1.03	1.24	1.78***	1.09	0.51*
	(0.200)	(0.478)	(0.290)	(0.454)	(0.205)
Black * MMM	1.13	1.01	4.94***	1.06	2.84**
	(0.314)	(0.530)	(1.406)	(0.562)	(1.320)
Hispanic * MMM	1.10		0.82		2.50
	(0.825)		(0.762)		(4.505)
Other * MMM			1.00		19.04*
			(0.767)		(29.686)
Constant	0.14***	0.06***	0.11***	0.05*	0.29

	(0.039)	(0.018)	(0.031)	(0.074)	(0.289)
N	2,984	3,102	2,595	309	983

Robust standard errors are reported below each coefficient in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Success. Consistent patterns emerge across all three general education mathematics courses targeting regular students. Female students in each course have a higher chance at succeeding than their male counterparts, and the higher a student's ACT score the higher the likelihood of succeeding in a course. We find that students who were retaking a course have significantly lower odds of succeeding compared to those taking a course for the first time. With respect to the racial/ethnic disparities, we see that under conventional instruction Black students have 38%, 29%, and 49% lower odds of succeeding than White students in Foundations, College Algebra, and Elementary Calculus, respectively. Other students have 79% higher odds than White students to succeed in Elementary Calculus.

The MMM teaching pedagogy is significantly effective in increasing the odds of succeeding in Elementary Calculus — students exposed to the MMM have 78% higher odds of succeeding than those in traditional Elementary Calculus. Furthermore, the large magnitude and significance of the interaction of teaching method and race illustrates a particular benefit of this teaching method for Black students. In Elementary Calculus, Black students instructed via the MMM have 779% (computed as 1.78*4.94 -1) higher odds of succeeding than Black students receiving conventional instruction.

Interpreting coefficients within a logistic regression is difficult since while the marginal effect on the logit, or on the odds ratio, is constant, the marginal effect on the probability depends on the other variables (more generally the marginal effect depends on the probability). As an example, we can compute the predicted probability for a male black student taking Elementary Calculus for the first time who obtained the average overall ACT Math score of 19. Using our model, such a student has 29.0% chance of succeeding when taught traditionally and a 78.2% chance of succeeding if receiving the MMM instruction. This computation concretely illustrates the effectiveness of the MMM on increasing student success in Elementary Calculus.

Columns 4 and 5 of Table 4 illustrate the succeed regression results for DSPM students only. As with the regular student population, female students have a higher chance of succeeding, as do students with higher ACT scores. We see that the MMM method is not statistically significant in increasing student success rates in Foundations and is only marginally significant in increasing success in College Algebra. Black DSPM students in College Algebra taught with MMM pedagogy, however, do see a benefit with respect to Black student taught conventionally. In fact, we find a 45% (computed as 0.51*2.84 -1) increase in odds for success for Blacks taught College Algebra with the MMM.

Table 5. *Logistic Regression of Dropout Against Explanatory Variables*

	Regular			DSPM	
Variables	Foundations	Algebra	Calculus	Foundations	Algebra
Female	0.89	0.80*	0.72***	0.68	0.84

	(0.114)	(0.093)	(0.084)	(0.289)	(0.190)
ACT Math Score	0.93***	0.88***	0.93***	0.79*	0.92
	(0.018)	(0.018)	(0.016)	(0.103)	(0.079)
Redo	1.27	1.21	0.71**	1.75	1.02
	(0.207)	(0.186)	(0.110)	(1.097)	(0.378)
Black	0.90	0.69***	1.33**	0.79	1.98
	(0.125)	(0.093)	(0.177)	(0.439)	(1.354)
Hispanic	1.20	0.93	1.88		4.20
	(0.492)	(0.380)	(0.850)		(5.721)
Other	0.84	1.02	0.65		
	(0.403)	(0.314)	(0.235)		
MMM	0.49*	0.21	0.52**	1.55	1.19
	(0.186)	(0.212)	(0.136)	(0.966)	(0.772)
Black * MMM	1.78	5.09	0.40*	0.37	0.60
	(0.841)	(5.912)	(0.189)	(0.340)	(0.443)
Hispanic * MMM	1.65		1.31		
	(2.014)		(1.583)		
Other * MMM			4.33		
			(3.885)		
Constant	0.58	1.84	0.91	7.31	0.31
	(0.238)	(0.804)	(0.347)	(16.031)	(0.504)
N	2,984	3,102	2,595	305	959

Robust standard errors are reported below each coefficient in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Dropout. Table 5 shows that regular female students have a lower probability of dropping Calculus compared to male students. For example, female students have 28% lower odds of dropping out of Elementary Calculus compared to their male counterparts. We find a strong ACT score effect illustrating that students with higher ACT scores have lower odds of dropping out. This result holds across all courses thus illustrating the importance of background mathematical knowledge on retention. Students who are retaking a course are more likely to persist in Elementary Calculus and have 29% lower odds of dropping out.

As with the succeed variable, we compute predicted probabilities for a male black student obtaining the average ACT math score of 19 taking Elementary Calculus for the first time. Using our model, such a student has a 23.4% chance of dropping out if the student receives conventional instruction and a 6.0% chance of dropping out if the student receives MMM instruction.

Black/white differentials persist when comparing the probabilities of dropping out. Black students in College Algebra have 31% lower odds of dropping out compared to White students. The MMM is positive and significant for students taking Calculus. Calculus students in the MMM have about 48% lower odds of dropping out with respect to conventionally taught students. This positive finding provides evidence that the MMM is effective in increasing retention.

No significant relationships were found for the DSPM Algebra and Foundations courses in relation to drop out rates.

Discussion & Conclusion

Despite best efforts, hundreds of thousands of students are *not* succeeding in postsecondary general education mathematics courses each year. This situation is of particular concern, because failure to pass a required general education mathematics course jeopardizes one's ability to complete an undergraduate degree. In addition, this issue takes on an added dimension of urgency as the US struggles to improve both the overall percentage of citizens who attain a postsecondary degree as well as to close the educational attainment gap between minority and non-minority populations.

MMM was developed and implemented at UM with these factors in mind. Our results suggest that MMM was positive and significant for raising success rates particularly in Elementary Calculus. In addition, the results show that the MMM is a vehicle for closing the achievement gap between Black and White students. For the remedial DSPM courses, the MMM instructional method was not found to be statistically significant for retention; however, as seen in the descriptive Tables 1-3, students in the MMM courses showed a smaller number of overall dropouts. In general, our data suggest that the MMM increases success and decreases dropout rates. The positive results may be attributed to the structure and interactive nature of the MMM which forces a daily involvement on the part of the student. This type of active engagement along with the use of technology is in-line with reform pedagogy.

From a practical standpoint, postsecondary institutions need to find a cost effective, scalable, and impactful method to address low success rates in general education mathematics courses. After an initial start-up cost in establishing suitable computer labs, the MMM distributes department resources in a cost effective way. First, the MMM can employ undergraduate student assistants, rather than graduate students, second, because grading is automated in MyMathLab, this eliminates the need to have graders for these classes. This then frees up advanced graduate students to be employed as instructors instead of graders.

The MMM implementation has resulted in overall improved student success in Elementary Calculus, lower dropout rates in College Algebra, and lower costs. As such, future work is needed to perform a scientific, comparative evaluation of the model is in order to provide concrete statistical evidence of its validity and in turn offer motivation for scale-up.

Bibliography

Alldredge, J. R., & Brown, G. R. (2006). Association of Course Performance with Student Beliefs: An Analysis by Gender and Instructional Software Environment. *Statistics Education Research Journal , 5* (1), 64-77.

Anderson, J. R., Corbett, A. T., Koedinger, K. R. & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences, 4*(2), 167-207.

Barrow, L., Markman, L., & Rouse, C. E. (2009). Technology's Edge: The Educational Benefits of Computer-Aided Instruction. *American Economic Journal: Economic Policy , 1* (1), 52-74.

Bransford, J. & Schwartz, D., (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 3, 24, 61-100.

Brooks, D. W., (1997). *Web-teaching: a guide to designing interactive teaching for the World Wide Web*. New York: Plenum Press.

Caldwell, J. E. (2007). Clickers in the Large Classroom: Current Research and Best-Practice Tips. *CBE- Life Sciences Education , 6* (1), 9-20.

Chi, M., Siler, S., Jeong, H. & Hausmann, R., (2001). Learning from human tutoring. *Cognitive Science*, 25, 471-533.

Cretchley, P., Harman, C., Ellerton, N., & Fogarty, G. (2000). MATLAB in early undergraduate mathematics: An investigation into the effects of scientific software on learning. *Mathematics Education Research Journal , 12*, 219-233.

de La Beaujardiere, J. F., Cavallo, J., Hasler, A. F., Mitchell, H., O'Handley, C., Shiri, R. & White, R. (1997). The GLOBE Visualization Project: Using WWW in the Classroom. *Journal of Science,* 6, 1.

De Vega, M., Glenberg, A., & Graesser, A. (2008). *Symbols and Embodiment: Debates on Meaning and Cognition.* Oxford: Oxford University Press.

Fies, C., & Marshall, J. (2006). Classroom Response Systems: A Review of the Literature. *Journal of Science Education and Technology , 15* (1), 101-109.

Fletcher, J. D. (2003). Evidence for Learning From Technology-Assisted Instruction. In H. O'Neil, & R. S. Perez, *Technology Applications in Education: A Learning View* (pp. 79-100). Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.

Hauk, S., & Segalla, A. (2005). Student perceptions of the web-based homework program WeBWork in moderate enrollment college algebra classes. *Journal of Computers in Mathematics and Science Teaching , 24* (3), 229-253.

Heid, K. M., & Edwards, M. T. (2001). Computer Algebra Systems: Revolution or Retrofit for Today's Mathematics Classroom? *Theory Into Practice , 40* (2), 128-136.

Henningsen, M., & Stein, M. K. (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.

Hosmer, D.W., & Lemeshow, S. (2000). Applied Logistic Regression, New York: Wiley, ISBN 0471615536.

Judson, E., & Sawada, D. (2002). Learning from Past and Present: Electronic Response Systems in College Lecture Halls. *Journal of Computers in Mathematics and Science Teaching , 21* (2), 167-181.

Khan, B. H. (Ed.) (1997). *Web-based instruction*. Englewood Cliffs, N.J.: Educational

Technology Publications.

Koedinger, K. R., Anderson, J., Hadley, W. & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-43.

Lavicza, Z. (2009). Examining the use of Computer Algebra Systems in university-level mathematics teaching. *Journal of Computers in Mathematics and Science Teaching , 28* (2), 99-111.

Liao, Y.-k. C. (2007). Effects of Computer-Assisted Instruction on Students' Achievement in Taiwan: A meta-analysis. *Computers & Education , 48* (2), 216-233.

Lowther, D. L., Ross, S. M., & Morrison, G. M. (2003). When Each One Has One: The Influences on Teaching Strategies and Student Achievement of Using Laptops in the Classroom. *Educational Technology Research and Development , 51* (3), 23-44.

Matsuda, N. & VanLehn, K., (2005). Advanced Geometry Tutor: An intelligent tutor that teaches proof-writing with construction. In Proceedings of the 12th International Conference on Artificial Intelligence in Education.

Mayer, R. E. (2003). The Promise of Multimedia Learning: Using the Same Instructional Design Methods Across Different Media. *Learning and Instruction , 13* (2), 125-139.

Mayer, R. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist , 59* (1), 14-19.

Mayer, R.E., (2005). *Multimedia Learning*. Cambridge, MA: Cambridge University Press.

Means, B., Olson, K., & Singh, R. (1995). Beyond the Classroom: Restructuring Schools with Technology. *Phi Delta Kappan , 77* (1), 69-72.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School.* (J. D. Bransford, A. L. Brown, & R. R. Cocking, Eds.) Washington, D.C.: National Academy Press.

Nguyen, D. M., Hsieh, Y.-C., & Allen, G. D. (2006). The Impact of Web-Based Assessment and Practice on Students' Mathematics Learning Attitudes. *Journal of Computers in Mathematics and Science Teaching , 25* (3), 251-279.

O'Callaghan, B. R. (1998). Computer-Intensive Algebra and Students' Conceptual Knowledge of Functions. *Journal for Research in Mathematics Education , 29* (1), 21-40.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., et al. (2007). *Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004.* Washington, D.C.: National Center for Education Research.

Peschke, J. (2009). Moving Ahead to the Future By Going Back to the Past: Mathematics Education Online. *Journal of Computers in Mathematics and Science Teaching , 28* (2), 123-133.

Ringenberg, M. & VanLehn, K., (2006). *Scaffolding problem solving with annotated, worked-out examples to promote deep learning*. Lecture Notes in Computer Science, 4053. Springer Berlin/Heidelberg.

Rouet, JF., (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Erlbaum.

Sandholtz, J. H., Ringstaff, C., & Dwyer, D. C. (1997). *Teaching with Technology.* New York: Teachers College Press.

Schoenfeld, A. H. (1994). *Mathematical Thinking and Problem Solving.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Smith, R. S. (2007). Spreadsheets in the Mathematics Classroom. In K. H. Ko, & D. Arganbright (Ed.), *Enhancing University Mathematics: Proceedings of the First KAIST International Symposium on Teaching. CBMS Issues in Mathematics Education Volume 14.* Washington, D.C.: American Mathematical Society.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction for Advanced Knowledge Acquisition in Ill-Structured Domains. In T. M. Duffy, & D. H. Jonassen, *Constructivism and the Technology of Instruction: A Conversation* (pp. 57-76). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Sweller, J. & Chandler, P., (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185-233.

VanLehn, K., (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education,* 16, 3.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A. & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31, 3-62.

Vanlehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., et al. (2005). The Andes Physics Tutoring System: Lessons Learned. *International Journal of Artificial Intelligence in Education , 15* (3), 147-204.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction , 16* (1), 3-118.

Yaron, D., Cuadros, J. & Karabinos, M., (2005). Virtual Laboratories and Scenes to Support Chemistry Instruction. *About Invention and Impact: Building Excellence in Undergraduate STEM Education*, 177 - 182.

Zerr, R. (2007). A quantitative and qualitative analysis of the effectiveness of online homework in first-semester calculus. *Journal of Computers in Mathematics and Science Teaching , 26* (1), 55-73.

USING CONCRETE METAPHOR TO ENCAPSULATE ASPECTS OF THE DEFINITION OF SEQUENCE CONVERGENCE

Paul Christian Dawkins
Northern Illinois University
dawkins@math.niu.edu

This study traces how one real analysis student was able to develop a property-based rather than exemplar-based notion of sequence convergence by first instantiating conditions on the parameters of sequence convergence in a metaphorical domain. The analysis uses the radical constructivism framework of cognitive development (von Glasersfeld, 1995). Once he had encapsulated the conditions that related the parameters within the formal definition, he was able to accommodate the conditions into a mathematical schema in which he could reason flexibly about convergent sequences in ways compatible with standard formal practice. I observe his language use and the various lines of reasoning that his concept of the sequence convergence definition did or did not support over the course of a month to identify changes in his convergence schema.

Key words: Real analysis, concrete metaphor, sequence limit definition, radical constructivism

Mathematicians as a community uphold abstraction and generality as hallmarks of mathematical reasoning. The pursuit of abstraction and generality has guided the community toward lines of reasoning built upon definitions and properties rather than examples or categories of examples. The power of formal definitions and deduction thereupon lies in their ability to simultaneously represent entire classes of mathematical objects that satisfy the conditions of the definition. There is broad deductive power in the transition from “This sequence converges” to “For any convergent sequence…” However, research upon the transition from case-based and computational reasoning to definition-based and deductive reasoning (a subset of what has been called Advanced Mathematical Thinking) poses many obstacles for students (Alcock & Simpson, 2002; Tall, 1991). In the realm of calculus and analysis, conceptual aspects of limit have proven quite difficult (Ely, 2010; Oehrtman, 2009) prior to the difficult transition to formal limit reasoning based upon multiple-parameter, multiple-quantifier definitions (e.g. $\forall\, \varepsilon > 0\ \exists\, K \in N$ such that $\forall\, n \geq K, |a_n - L| < \varepsilon$). The present study investigates how a real analysis professor’s introduction of a concrete metaphor supported one student’s transition toward more formal reasoning both in his understanding and use of the definition of sequence convergence.

Relevant Literature

Three primary pathways have been described for the development of formal definition-based reasoning about definitions of limits: 1) exploring a range of examples to develop a definition that uniquely describes them (Swinyard, 2008), 2) developing "generic examples" through graphical explorations (Pinto & Tall, 2002), and 3) extracting meaning from the formal definition itself (Dubinsky, Elterman, & Gong, 1988). The first pathway via capturing examples holds difficulty in that it takes time and several studies have noted significant differences among the ways different students reason about mathematical definitions (Alcock & Simpson, 2002;

Vinner, 1991) and differences between the ways that students and mathematicians treat definitions (Edwards & Ward, 2008). Alcock and Simpson (2002) distinguished between those students who used examples or sets of examples with which they were familiar to reason about novel examples and those students who used the formal definition to reason about novel examples. Though neither form of reasoning are absent in standard mathematical practice, the mathematical community values the latter because any reasoning depending upon the definition necessarily applies to the entire category of objects satisfying the definition. Edwards and Ward (2008) similarly distinguished between definitions that establish their associated category (stipulated definitions) and definitions that intend to capture a pre-existing category (extracted definitions). In proof-based settings, mathematicians treat definitions as stipulated while students tend to treat definitions as extracted, and this distinction strongly determines the lines of reasoning and conclusions that an individual develops. Vinner (1991) described a similar distinction between students reasoning primarily with their concept image (the set of all ideas, images, examples, and processes associated with a given definition) and students reasoning primarily with the concept definition (the standard set of words used to define a given concept).

Other research has noted that students are capable of developing rich graphical representations for reasoning flexibly about sequences and limits, but in some cases this proves to support development of more formal reasoning in terms of the definition (Pinto & Tall, 2002) and in others it seems to inhibit that development due to the convincing power of the graphical reasoning (Alcock & Simpson, 2004). Finally, in many classrooms in which definitions are presented in final form, students are asked to develop understanding and fluency with the formal definition simply by using it to verify convergence of particular examples or to prove theorems about classes of limits. The present study describes an alternate pathway that one student traversed toward fluency with the standard formal definition using a concrete metaphor for convergence the professor developed during instruction.

Oehrtman (2009) described the metaphors calculus students used to reason about limits contextually and computationally. However, the metaphors he described were not metaphors for the formal definition, but rather metaphors for the notion of limit itself such as numerical approximation or collapsing a dimension. Current work is also exploring how students successfully used sets of examples and the approximation model of limits to develop the formal definition of sequence limits, series limits, and point-wise convergence of a Taylor series (Martin, Oehrtman, Roh, Swinyard, Hart-Webber, current conference proceedings). This study, being informed by Swinyard and Oehrtman's previous work, extends the body of Realistic Mathematics Education (Freudenthal, 1973, 1991) that takes the instructional stance that students should be guided to rediscover mathematics whenever possible. Rediscovery need not follow the historical process by which mathematics was developed, but rather seeks to create pathways by which students can viably develop their own mathematical constructions that approximate standardized algorithms, definitions, theorems, and proofs.

Theoretical Framework

My theoretical framework draws primarily from radical constructivism as articulated by von Glasersfeld (1995). Radical constructivism posits that concepts do not necessarily reflect the objective nature of external reality, but rather are tools that minds use to organize and deal with ongoing experiences. The viability of conceptual structures lies in their ability to account for new experiences and guide the individual's actions toward desired results, often negotiating problematic situations. In the context of mathematics, individuals construct mathematical conceptions to organize their reasoning and to satisfy the demands of the mathematical tasks in which they are engaged. For a mathematician, these tasks may be self-imposed and be guided by a cultural set of standards for the acceptability of posing and "solving" tasks. For a student, tasks often appear primarily as an element of their educational experience such that tasks are usually introduced by the instructor and standards of "solving" a task derive from the acceptability of their answers in the eyes of their classmates and instructor. In any case, students develop conceptual structures to address the demands of mathematical tasks and assess the viability of those structures either according to an internal sense of coherence (in terms of their pre-existing conceptions) or according to external feedback their mathematical interactions with others elicit.

Once a student has developed conceptual structures for dealing with various types of tasks, there are two primary processes by which new experiences or tasks interact with that structure, which is referred to as a schema. When a student accesses a schema to reason about or guide their action about a novel task or experience, they are *assimilating* the experience into that schema. For instance, a student may have never seen the number 240,831, but they may be able to ascertain many aspects of the quantity because they recognize that this set of numerals can be understood using the place-value number schema. The student may assimilate like this automatically trusting in their line of reasoning until they receive some feedback that their number schema has not led them to a coherent result or they receive input that can no longer be easily assimilated: for instance when they are told that .999… = 1. In cases where either a sense of internal consistency, feedback from teachers or classmates upon their own thinking, or new ideas that are presented fail to be viably assimilated by their existing schema (which is referred to as *disequilibrium*), students may alter the schema or reorganize their assessment of the novel input which is called *accommodation*. In the case of .999… = 1, the student must engage both processes: first they must identify that their place-value number schema applies to the numbers (assimilation), but they must alter that schema to allow for a single quantity to have multiple representations (accommodation).

The process of organizing sets of exemplars into categories to which a single schema applies has the effect of unitizing or chunking thoughts to ease future processing. Most adults do not spend time reflecting on or adjusting their number schema because it is able to quickly and sufficiently assimilate all of their ongoing experiences with numbers. In this way, people refer to "number" without perturbation, because this abstract set of objects with it's associated rules, properties, and operations can be thought of as a single conceptual structure. Unitizing what at earlier stages of development was a myriad of associated examples, concepts, and processes is called *encapsulation*.

Methods

This study was conducted at a mid-sized university (25,000 students) in the USA. All data in this study was gathered during one 15-week semester of undergraduate real analysis taught by a research mathematician. At this university, this course generally includes a proof-based development of real numbers, sequences, limits of functions, and continuity. At the time of data gathering, the professor had been teaching for over 10 years and had previously taught undergraduate analysis thrice. She has received multiple teaching awards and is widely regarded as an excellent, if not difficult, instructor. The class met twice weekly for 80 minutes at a time and began with 23 students.

Data gathered includes field notes of class meetings, biweekly professor interviews, weekly student interviews with a small group of volunteers from the class, copies of students' written class notes, and copies of interview participant exams. Written field notes recorded all communication written on the blackboards, major aspects and key quotes from professor and student verbal communication, and physical gestures displayed in the discussion of course material.

All interviews were audio-recorded and any written records of the interactions maintained. The student interview participants were selected from among mathematics majors who volunteered so as to represent a variety of final grades in the "Intro to Proofs" course that serves as a pre-requisite to analysis at this university. Thus the 6 interview participants included one or two students each who made an A, a B, and a C in "Intro to Proofs."

The interviews (at various times) invited students to:

- recall and explain definitions, theorems, and proofs,
- explain and assess mathematical statements and proofs,
- recall and explain aspects of classroom discussion,
- relate the nature, content, extent, and quality of their group interactions outside of class,
- complete homework activities or other novel activities,
- explain their reasoning on written exam questions,
- articulate their confusion about any mathematical topics with which they were uncomfortable, and
- comment about their course experiences both positive and negative.

Interviews ascertained students' recall and interpretations of classroom explanations and discussions as well as their strategies and success on questions and activities presented to them.

The rather general nature of the lines of questioning reflects the research intention that interviews reflect and respond to particular events and elements of the classroom dialogue. Interview questions were developed throughout the semester to pursue student understanding and perception of things such as class discussions, particular diagrams, or chains of reasoning that were expressed during class meetings. The interview questioning often moved from very open-ended to very particular so as to initially allow interview participants flexibility in their choices of verbalization and representation. Then, if needed, interview questions centered participants' focus onto particular elements from the classroom presentation that were of interest to the professor or appeared to the researcher central to the fidelity of classroom communication. For example, the interview protocol may begin by asking a student, "What does it mean for a sequence to converge to a point?" Later, students would be asked about particular aspects of the definition or the explanation thereof such as, "What does 'arbitrary but fixed' mean and why did [the professor] discuss it?" or "In the definition of convergence, what are ε, K, and n?" Finally, a student might be asked to construct a proof that a particular sequence converges to a given limit.

Interviews were transcribed and then coded according to the open coding method described by Strauss and Corbin (1998). Categories varied in nature. Some captured broad sets of interactions such as examples of definition use or comments on the structure of mathematics. Other categories organized every appearance of a given question or example during an interview.

Results

Beginning early in the semester, the professor discussed definitions in two distinct contexts: the concrete metaphorical domain and mathematical domain. The class discussed their reasoning within these two domains in terms of three distinct linguistic registers (sets of words associated by meaning or use), which we shall call the metaphorical register, the intuitive register, and the formal-symbolic register. In the context of the definition of sequence convergence, the professor presented the class with an intuitive register definition meant to represent common notions of convergence that said, "A sequence (a_n) converges to L if a_n gets closer and closer to L as n gets larger and larger." However, she pointed out that in the case of $(4 - {}^{1}/_{n})$, five is a possible limit of this sequence by this definition. Figure 1 shows the number line representation the professor used to display this sequence; she commonly used such number line diagrams to discuss sequences. As students proposed different ways to amend the intuitive register definition to be more specific, Zell told the teacher that "if you want to find a party, see where everyone is at." The professor agreed that they needed to find "where everyone is at," and clearly no one is at 5.

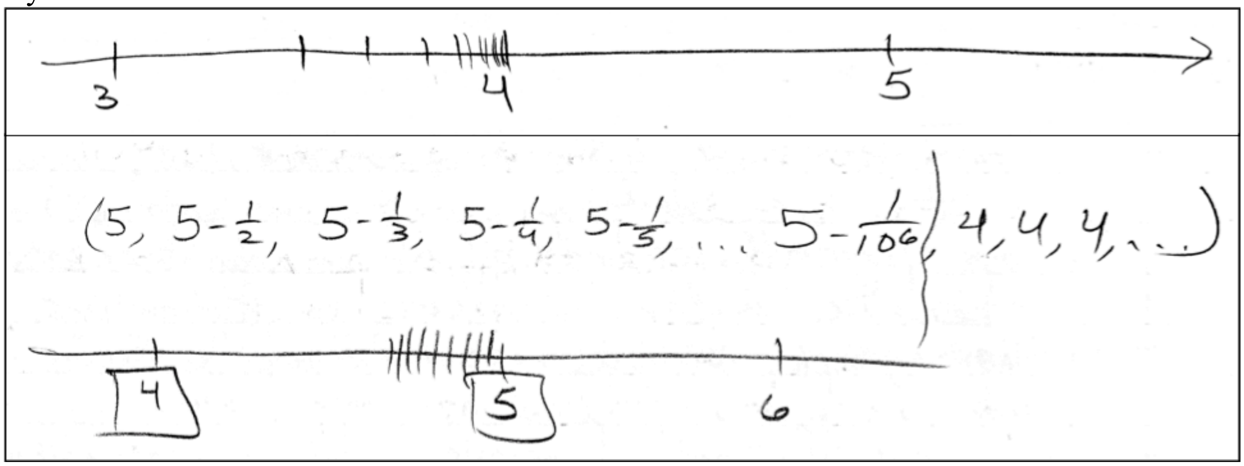


Figure 1. The professor's number line and term-enumeration representations of sequences.

The professor then extended the metaphor asking, "How many terms make a party?" She provided another example that was defined as $5 - {}^{1}/_{n}$ for the first million terms and 4 for every latter term. Figure 1 displays the term-enumeration and number line representations that the professor used to display this sequence. She posed the question, "How many people have to be at the party," to which Locke responded "infinitely many." The professor invited the students to talk for a while and share their ideas about how they should form their formal definition. To bring the discussion together again, she wrote on the board a revised intuitive register definition that stated,

"$*_1$ A sequence converges to the real number L if we can make the terms of the sequence stay as close to L as we wish by going far enough out in the sequence."

She verbally mapped this statement into the concrete metaphorical domain saying it is only a party if for any size party you pick, after some point everyone shows up at the party. She then began to ask the class how they could make various aspects of this informal definition "more rigorous." They replaced the previous definition with the following:

"$*_2$ (x_n) converges to L if, given $\varepsilon > 0$, all terms after a certain term x_K are in $(L - \varepsilon, L + \varepsilon)$."

In another formulation, she stated verbally that, "only finitely many guys can be outside the room for you to have a party." The process of translation between the registers ended on the definition:

"* A sequence (x_n) converges to the real # L if given $\varepsilon > 0$, there exists $K \in N$ such that $\forall n \geq K,\ x_n \in (L - \varepsilon, L + \varepsilon)$."

During later lectures, the professor began to refer to these terms outside the party as "stragglers." Immediately after completing the definition of sequence convergence, the class worked to prove that $3n/_{2n+1}$ converges to 1.5, and they translated $3n/_{2n+1} \in (1.5-\varepsilon, 1.5+\varepsilon)$ into absolute value notation to say $\left|3n/_{2n+1} - 1.5\right| < \varepsilon$ in line with their work several weeks earlier on epsilon neighborhoods. Their proof, once completed, only justified the claim in terms of absolute value notation.

Thus the professor translated the intuitive register definition into the formal-symbolic register by replacing "stay as close to *L* as we wish" with epsilon neighborhood conditions and "going as far out as we wish" with index terminology. However, for the next several weeks of class meetings that covered the section on sequences, the professor continued to refer verbally to the party metaphor though everything written on the board was either graphical, intuitive register, or formal-symbolic register.

Linguistic Register	Sequence Elements	Proximity	Sequence order	"Cut off point"
Concrete metaphorical	People	Location	Time	Eventually
Informal mathematical	Terms	Closeness	Physical Arrangement	"Far enough out in the sequence"
Formal-symbolic	a_{n}	ε	Indices (n)	K

Table 1. Sequence convergence in the various registers.

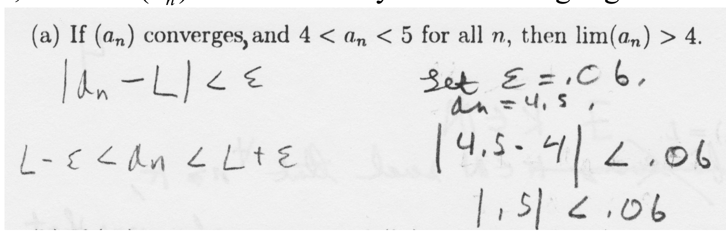
Tidus emerged as an instructive case of reasoning about sequence limits because of the central role the party metaphor played in his reasoning about limits. Over the course of talking with Tidus during interviews, several key aspects of his concept image of sequence convergence repeatedly appeared. When possible, each primary element of his concept image will be connected with tasks he attempted during exams or interviews to understand what forms of reasoning his concept image and concept definition supported and failed to support. Since all of the four interviews and two tests that dealt with aspects of sequence convergence fell within a calendar month, any quotes will be marked with their source type and day of the month: I-2, T-8, I-17, I-23, I-28, or T-29.

Neighborhood as place

Tidus consistently referred to the epsilon neighborhood as a place or a set rather than a measure of closeness between sequence terms and the limit value. He expressed his understanding of the definition of sequence convergence saying, "at some point … eventually all of those numbers will be in four's neighborhood" (I-2). He articulated that same definition later saying, "For any epsilon that you pick, an infinite amount of terms will be in that epsilon neighborhood and a finite amount of terms will be outside" (I-17). This appeared on the first test (T-8) in that six different times he juxtaposed absolute value and set-type expression such as $|x_n - 5| < .01/_2$ alongside $x_n \in \left(5 - .01/_2, 5 + .01/_2\right)$ or $|a_n - L| < \varepsilon$ alongside $L - \varepsilon < a_n < L + \varepsilon$. Consistently thereafter when discussing sequence convergence, Tidus would talk about numbers or terms being "in the epsilon neighborhood" rather than using closeness or approximation language. The only exception to this in the context of sequence convergence appeared on the first exam when

Tidus argued that $(a_n) \to 0$ and (b_n) converges $\Rightarrow (a_n b_n) \to 0$ by noting that "a_n gets closer and closer to zero."

This physical location view on epsilon neighborhoods appeared to limit Tidus' ability to reason about the artibrarity of epsilon or to change his value of epsilon. In the same way that the size of a house in the party metaphor cannot change, Tidus rarely referred to multiple values of epsilon in the same problem early on. This viewpoint led him to assert that "If (a_n) converges and $4 < a_n < 5$ for all n, then $\lim(a_n) > 4$" is false by the following argument:


(a) If (a_n) converges, and $4 < a_n < 5$ for all n, then $\lim(a_n) > 4$.

$|a_n - L| < \varepsilon$

set $\varepsilon = .06$, $a_n = 4.5$,

$L - \varepsilon < a_n < L + \varepsilon$

$|4.5 - 4| < .06$

$|.5| < .06$

Figure 2. Tidus' limit argument (T-8).

Tidus concluded "$\lim(a_n) > 4$ and a_n is trapped when $a_n = 4.5$ and $L = 4$" (T-8). Though his example and argument do not match the formal definition, it displays how his understanding of convergence at this point involved a single value for epsilon and "trapping" all of the values of the sequence in that epsilon neighborhood.

Tidus produced a mostly accurate proof that if a sequence converges to 5, then it cannot converge to 5.01. His chose $\varepsilon = .01/2$ and pointed out that the epsilon neighborhoods were disjoint. Since it is impossible for the terms of the sequence to be in the two neighborhoods at the same time, this produced a contradiction (T-8). Thus, Tidus' image of the epsilon neighborhood as a place supported success with certain types of reasoning, and success on tasks in which the fact that numbers are either inside or outside of the neighborhood proved important.

"Finding the K" as convergence

For some time, Tidus did not seem to reason at all about changing the value of epsilon relative to the same limit, but literally would "pick any" and leave it static thereafter. He understood that he could pick a convenient epsilon and did so on several tasks on the first test. During the second interview (I-17), he said:

- "If you show that the [sequence] can't go out [of the epsilon neighborhood], then it's going to converge."
- "It has to converge to an L in any epsilon neighborhood."
- "There should be one number in the [sequence] where the [sequence] converges to that point, where all of the terms are in that epsilon neighborhood."

Though Tidus does mention that the condition must hold "for any epsilon neighborhood," he also speaks in the first and third quotes as if convergence to the limit is not a global property of the sequence, but it is something that the sequence is "going to do" relative to the "one number in the [sequence] where the [sequence] converges to that point." Similarly during the third interview (I-23) he said, "We are finding a K_1 where it is going to converge to 5." Thus the focus of Tidus' view of convergence was on the bounding of infinitely many terms after K rather than upon the size of epsilon or the universal quantifier associated.

This notion supported his reasoning on some types of tasks. On the first test (T-8), the professor included a true-false question that stated: "Suppose the sequence X converges to -1. For $K \in N$, let X^K denote the sequence obtained by changing the first K terms of X to the value K

and leaving the rest of the terms as they are. Then, for all $K \in N$, the sequence X^K converges to -1." The professor spent some time during class articulating the point that a finite number of terms have no effect on the convergence of a sequence. Tidus' image of sequence convergence that focused on the existence of a point after which the terms must stay in the neighborhood allowed him to dismiss the changing of *K* terms as irrelevant for convergence. After writing out an example sequence *X* and X^3, he argued that the statement is true because, "If you change the value of the stragglers and did not the value of the terms [in] the *ε*-neighborhood, the sequence will still converge to *L*" (T-8). He used party metaphor language to refer to the finitely-many terms outside the neighborhood, and possibly this supported his understanding that their absence did not affect "the party."

The emergence of the arbitrariness of epsilon

During the third interview, Tidus began to juxtapose his use of epsilon language with "closeness" language in the context of the definition of Cauchy sequences. In the course of arguing why Cauchy sequences must be bounded, he said, "Well if they are getting close to each other, at some point they are going to be in each other's, I guess you could say epsilon neighborhood? Something like that because if they are getting closer to each other they can't be going out anywhere else" (I-23). Tidus continued to use his informal domain expression "at some point," and "in each other's… epsilon neighborhood," but he used epsilon as a means of expressing proximity as well.

Toward the end of the third interview, he wrote down a form of the formal-symbolic definition that began, "Let $\varepsilon > 0$…" and then appended to the end of his explanation, "and it should work for every epsilon" (I-23). This was the first time that Tidus has specifically referenced the generality or arbitrariness of epsilon in conjunction with the formal definition. When probed about the difference between "Given $\varepsilon > 0$…" and "For every $\varepsilon > 0$…", Tidus said "I don't see how it could be different… 'Cause either way you are going to have to pick… an epsilon. For every you are saying you are taking any epsilon, that *K* is going to be different for every epsilon… For a different epsilon there will exist a different *K*" (I-23). This was also the first time during interviews that Tidus expressed any awareness of this "dependence" between the quantities.

During Tidus' final interview regarding sequences (I-28), he worked on review problems for his upcoming exam and considered the question,

" (a) If (x_n) is a sequence and x_K is a term of the sequence such that $|x_n - x_K| < 0.1$ for all $n \geq K$, then which of the following, if any, are correct?

i) (x_n) is bounded
ii) (x_n) is Cauchy
iii) (x_n) has a Cauchy subsequence

(b) Let (x_n) be a sequence such that for every $\varepsilon > 0$, $|x_n - x_m| < \varepsilon$ for all $m,n \in N$. Then which of the following, if any, are correct?

i) (x_n) is Cauchy
ii) (x_n) is a constant sequence"

Tidus originally indicated that the sequence in part (a) must be convergent and thus Cauchy, but then caught himself and said, "Oh, but it's not for every epsilon… there could be an epsilon where it doesn't work" (I-28). About the latter question, he noted, "It's not picking a certain term where this happens… that could only happen if it's a constant sequence… cause it doesn't matter what term in the constant sequence you pick for it to happen, it's going to happen everywhere" (I-28). When pressed to prove why every term of the second sequence must be equal, Tidus could not produce the proper argument, but his image of what must happen if there was not a "certain term" led him to believe the sequence must be constant. He argued, "because here it doesn't say… that you have to pick a certain term in the sequence where this happens… [which] like the convergent ones [allows] stragglers." In both cases, the tasks required Tidus to reason about the definition of convergence or the definition of Cauchy in terms of multiple values of epsilon at the same time, and he was successful in reaching accurate conclusions.

Discussion

Tidus' initial reasoning about limit in which he never considered multiple values for epsilon showed the central place that the party metaphor took in his concept image of sequence convergence. His argument that the constant sequence $a_n = 4.5$ converges to 4 when $\varepsilon = .6$ conflicts with his calculus-based knowledge of sequence convergence. Tidus readily displayed in other contexts that he knew constant sequences converged to the constant value. When he made this argument, he reasoned about sequence convergence in terms of his concept image that was dominated by the party metaphor. For a neighborhood of size 0.6, all of the terms a_n are in 4's neighborhood, to use his parlance.

Tidus' reasoned and spoke about epsilon neighborhoods as places rather than treating epsilon as a measure of proximity. This further supports the assertion that Tidus' evoked concept image was dominated by properties of concrete space and physical location rather than properties of the number line such as quantity and numerical difference. Thus, when Tidus used terminology from either the concrete metaphorical or mathematical registers, these terms were most directly referencing elements of his concept image, and particularly the party metaphorical image of sequence convergence. This is true even for his argument on the first exam where he makes no specific reference to the party metaphor, but he forms an argument that directly conflicts with his computational knowledge of sequence convergence.

The professor articulated the conditions of sequence convergence in different linguistic registers; in particular, she described the interrelationship between the parameters ε, K, and n in the metaphor of a party and in the mathematical realm of sequences, neighborhoods, and indices. Because Tidus's schema for people attending a party was more accessible and well-developed than his mathematical schema for interpreting logical dependencies, he instantiated his concept of the definition of sequence convergence using his experiential schema of people and buildings. He was able to quickly and easily establish the interdependencies between a given epsilon, the cut-off point K, and the people at the party as represented by n. However, because the party scheme did not easily assimilate the universal quantifier on the size of the building (epsilon), it was not adopted into his conception of sequence convergence. Hence, his intermediate definition reflected this set of interrelationships between the parameter quantities: "at some point… eventually all of those numbers will be in four's neighborhood" (I-2).

Tidus received affirming feedback upon this definition schema because it properly facilitated his reasoning on certain tasks such as proving that the limit of a sequence is unique or that changing finitely many terms of a convergent sequence does not affect convergence. As he used this schema to solve new tasks and assimilate further experiences with sequences, the interrelationships between a given epsilon, the found K, and all n beyond K became encapsulated into a single concept in Tidus' thinking. This represented itself in his compact reference to the entire condition (his party metaphor-based ε-K-n schema) as convergence when he said, "We are finding a K_1 where *it is going to converge* to 5" (I-23). It was during this third interview, once Tidus displayed evidence of having encapsulated his "finding the K" such that… condition, that he also showed signs of assimilating the universal quantifier onto epsilon within this schema. I shall refer to his encapsulated schema as Tidus' K-condition. After writing the formal concept definition, which he previously reasoned about using his K-condition, he added, "and *it should work* for every epsilon" (I-23).

It is not evident from the interview data what exactly caused the disequilibrium that guided Tidus to begin this process of assimilation. He admitted during the second interview (I-17) that even though he could reproduce proofs that a given sequence converges to a real number L, he did not understand why the proof satisfied the definition (which he stated in terms of his K-condition). He lost points on his exam for the way in which he wrote the formal definition (T-8). He began to assimilate new ideas into his ε schema when he used it to describe the proximity of sequence terms in the Cauchy definition. He thus experienced both internal and external feedback that would have made him reflect upon the complete viability of his K-condition as the definition of sequence convergence.

Tidus was now able to reason more fluidly about the complete concept definition because it had been reduced to, "For every $\varepsilon > 0$, the K-condition holds." Encapsulating the K-condition allowed Tidus to reason more freely without having to actively coordinate the roles and quantifiers attached to K and n. By the final interview, Tidus displayed reasoning about sequence convergence that was compatible with standard interpretations and uses of that definition. He leveraged his K-condition schema to reason about a novel definition that lacked a K value. Tidus still used terminology that arose from the party metaphor when he identified that the absence of K doesn't allow any "stragglers." However, his reasoning did not display the limitations of the party metaphor because he reasoned using multiple possible values of epsilon. In this way, he had now accommodated the K-condition into a broader mathematical schema of the sequence convergence definition. This schema that organized the roles and quantifiers on the parameters ε, K, and n was now viable for assimilating novel definitions presented in the formal-symbolic register. The use of the term "stragglers" may even reflect the lack of a comparable term in the mathematical register; the phrase "tail of the sequence" is often used for the set of sequence terms after the K^{th} term, but no standardized term refers to those terms prior to the K^{th} term. In this way, Tidus' schema that had accommodated the K-condition from the party metaphor contained more tools for reasoning and communication than the formal statement of the definition entailed on its own.

Tidus' transition from the K-condition to reasoning in terms of the standardized definition stands in contrast to Alcock and Simpson's (2002) notion of inverting the property-category relationship in the sense that he did not begin by reasoning with examples or sets of examples, but rather developed a concrete mental model in which to conceptualize the logical conditions of sequence convergence. Tidus shifted the properties that for him characterized sequence convergence by abstracting from the concrete domain to the mathematical domain in which multiple values of epsilon can (and often must) be considered. This pathway toward the formal definition of limit did not begin with sets of examples, graphical representations, or the formal definition, but rather arose from a concrete metaphorical domain. Though Tidus could reproduce the statement of the concept definition of sequence convergence from memory, he interpreted that statement in terms of the K-condition couched in the party metaphorical domain until the point at which he encapsulated the *K*-condition and was able to accommodate the universal quantifier for epsilon.

Though it introduced some false lines of reasoning, the party metaphor proved formative for Tidus' reasoning because it allowed him to develop a property-based conception of convergence rather than a computation-based or exemplar-based conception. Tidus' intermediate conceptions of sequence convergence supported some lines of reasoning well but not others. By the end of the period of instruction, Tidus displayed the ability to accurately interpret novel conditions on sequence behavior stated in the formal-symbolic register. However, his reasoning still reflected useful aspects from earlier stages of his concept development. Thus the party metaphor appeared a non-trivial stepping-stone in Tidus' movement toward a concept definition and concept image of sequence convergence compatible with standard interpretations and toward fluency in the formal-symbolic register.

References

Alcock, L. & Simpson, A. (2002). Definitions: dealing with categories mathematically. *For the Learning of Mathematics, 22*(2), 28-34.

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.

Dubinsky, E., Elterman, F., & Gong, C. (1988). The student's construction of quantification. *For the Learning of Mathematics*, *8*(2), 44-51.

Edwards, B. & Ward, M. (2008) The Role of Mathematical Definitions in Mathematics and in Undergraduate Mathematics Courses. In Carlson, M. & Rasmussen, C. (Eds.) *Making the Connection: Research and Teaching in Undergraduate Mathematics Education MAA Notes #73* (pp. 223-232). Washington, DC: Mathematics Association of America.

Freudenthal, H. (1973). *Mathematics as an Educational Task*. Dordrecht: D. Reidel Publishing.

Freudenthal, H. (1991). *Revisiting mathematics education: The China lectures*. Dordrecht: Kluwer.

Oehrtman, M. (2009). Collapsing Dimensions, Physical Limitation, and other Student Metaphors for Limit Concepts. *Journal for Research in Mathematics Education, 40*(4), 396-426.

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.

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

Swinyard, C. (2008). *Students' Reasoning about the Concept of Limit in the Context of Reinventing the Formal Definition*. Doctoral Dissertation, Portland State University, Portland, OR.

Tall, D. (1991). *Advanced Mathematical Thinking.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In Tall, D. (Ed.) *Advanced Mathematical Thinking* (pp. 65-80). Dordrecht, The Netherlands: Kluwer Academic Publishers.

von Glasersfeld, E. (1995). *Radical constructivism: a way of knowing and learning.* London, England: Falmer Press.

Concepts Fundamental to an Applicable Understanding of Calculus

Leann Ferguson and Richard Lesh
Indiana University, Bloomington
leafergu@indiana.edu and ralesh@indiana.edu

Calculus is an important tool for building mathematical models of the world around us and is thus used in a variety of disciplines, such as physics and engineering. These disciplines rely on calculus courses to provide the mathematical foundation needed for success in their discipline courses. Unfortunately, many students leave calculus with an exceptionally primitive understanding and are ill-prepared for discipline courses. This study has begun to work with presumed experts (undergraduate mathematics and other discipline faculty members) to develop a small number of prototype tasks that will elicit, document, and measure students' understanding of a few calculus concepts the faculty participants believe to be essential to successful academic pursuits within their respective disciplines. This paper discusses the data and analysis from the first round of interviews. Implications of these findings for calculus curriculum are presented.

Keywords: Calculus, understanding, model-eliciting activities, design research

According to Ganter and Barker (2004):

> *Mathematics can and should play an important role in the education of undergraduate students. In fact, few educators would dispute that students who can think mathematically and reason through problems are better able to face the challenges of careers in other disciplines—including those in non-scientific areas. Add to these skills the appropriate use of technology, the ability to model complex situations, and an understanding and appreciation of the specific mathematics appropriate to their chosen fields, and students are then equipped with powerful tools for the future.*
>
> *Unfortunately, many mathematics courses are not successful in achieving these goals. Students do not see the connections between mathematics and their chosen disciplines; instead, they leave mathematics courses with a set of skills that they are unable to apply in non-routine settings and whose importance to their future careers is not appreciated. Indeed, the mathematics many students are taught often is not the most relevant to their chosen fields. For these reasons, faculty members outside mathematics often perceive the mathematics community as uninterested in the needs of non-mathematics majors, especially those in introductory courses.*
>
> *The mathematics community ignores this situation at its own peril since approximately 95% of the students in first-year mathematics courses go on to major in other disciplines. The challenge, therefore, is to provide mathematical experiences that are true to the spirit of mathematics yet also relevant to students' futures in other fields. The question then is not whether they need mathematics, but what mathematics is needed and in what context. (p. 1)*

These claims of Ganter and Barker detail the rationale for The Mathematical Association of America's (MAA) Curriculum Foundations Project (CFP, http://www.maa.org/cupm/crafty/cf_project.html). The CFP studied the first two years of undergraduate mathematics curriculum. Portions of the mathematics community and its partners, what I refer to as "client" disciplines (e.g., biology, business, chemistry, computer science, several areas of engineering), worked together to generate a set of recommendations that have assisted mathematics departments plan their programs to better serve the needs of its client disciplines (Ferrini-Mundy & Gücler, 2009).

The push to better serve the needs of client disciplines stemmed from the calculus reform efforts. Between the mid 1980s and the early 1990s, the undergraduate mathematics community engaged in a concentrated effort to overhaul the teaching and curriculum of beginning calculus (Ferrini-Mundy & Gücler, 2009). The heart of the reform was the concern over the depth and breadth of students' understanding of calculus (Douglas, 1986). This lack of understanding became especially apparent when students were asked to apply calculus in unfamiliar situations (Hughes Hallett, 2000), such as those encountered in client courses.

As Ganter and Barker (2004) implied, client department faculty often complain that students are unable to apply calculus in the client coursework. This coursework often asks students to use the calculus concepts in ways not familiar to them. For example, the minimization of average cost is done symbolically in calculus, whereas it is usually done graphically in economics (Lovell, 2004). At other times, even when the concept is used in a similar fashion, differences in notation or a lack of familiar cues (e.g., "maximum" or "minimum" in an optimization problem) derails students. Such difficulties in transferring knowledge between disciplines are stark indicators of a lack of understanding (Hughes Hallett, 2000). Thus, the reform called for fundamental changes in the curriculum and pedagogy of beginning calculus.

Therefore, the purpose of this study is to work with presumed experts to develop a small number of *prototype* tasks to measure understanding of a small number of calculus concepts that all agree are important. Correspondingly, three products will emerge from this study:

1. A determination of the fundamental calculus concepts.
2. An analysis of what it means to understand each concept.
3. A pool of tasks that elicit, document, and measure the type and level of understanding a student has of these concepts.

Together with the help of the researchers, the presumed experts have begun developing prototype tasks (product 3) that are informing products 1 and 2. This study is ongoing, and as such, the results discussed here represent the first iteration: the initial determination of the fundamental calculus concepts and what it means to understand a concept, as well as examples of potential tasks that elicit understanding.

Study Description and Methodology

The changes that took place during the reform years placed greater emphasis on conceptual understanding (Hughes Hallett, 2000) rather than procedural skills[1] (Ferrini-Mundy & Gücler,

[111] Ferrini-Mundy and Gücler did not define either *conceptual understanding* or *procedural skills*. To establish a common definition, for the purposes of this study, I define "understanding calculus beyond computations" as an intertwining of computational skill and conceptual understanding. It is the relationships between skill and understanding that give a student the power of application in a wide variety of settings. "In order to learn skills so they are remembered, can be applied when they are needed, and can be adjusted to solve new problems, they must be learned with understanding" (Hiebert et al., 1997, p. 6). Power and flexibility in

2009), but as Ganter and Barker (2004) pointed out, it has not been enough and the disconnect between what the client disciplines need and what the calculus courses provide still exists. Not only does the "what mathematics is needed and in what context?" question remain, it must be extended to include understanding and assessment, rather than merely calculus computations and procedures. Following in the footsteps of the MAA's Curriculum Foundations Project, while narrowing the mathematical content focus to calculus, this study began exploring the potential disconnect between the calculus taught in the mathematics classrooms and the calculus needed outside the mathematics classrooms at a particular undergraduate engineering institution. Through exploring the disconnect, this study began to identify some of the fundamental calculus concepts students need for successful academic pursuits outside the calculus classroom, to illustrate what it means to understand a concept, and to draft tasks that elicit, document, and measure student understanding of these concepts.

Describing the fundamental calculus concepts and creating the pool of tasks constitutes a "design research" study (Brown, 1992; Collins, 1992). In design research, the goal is to put people with different perspectives into situations that require them to express not only how they think about a concept, but to express it in a way that requires them to test and revise their way of thinking (Lesh, 2002). As such, each cycle, or interview session, includes divergent ways of thinking, selection criteria for the most useful ways of thinking, and sufficient means of carrying forward the ways of thinking so they may be tested and revised during the next cycle. Diversity, selection, and accumulation are necessary for iterative revisions to be passed forward.

The framework for this study can be thought of as a multi-tier design experiment (Lesh & Kelly, 2000). There are three tiers in this research project: 1) researchers/facilitators, 2) faculty members/researchers, and 3) students. For the type of experiment described here, the goal is not to produce generalizations about students, faculty members, or groups. Instead the primary goal is to work with presumed "experts" (instructors that have taught a course of interest two or more times) to develop a small number of prototype tasks that measure calculus understanding. Our role as researchers/facilitators is to initiate, document, and analyze the descriptions of the fundamental calculus concepts, frameworks for understanding these concepts, and associated tasks that were elicited from the faculty participants. The role of the faculty participants is to express, test, and revise their thinking about the concepts, frameworks, and tasks through lens of their teaching experience and student work. The third tier, students, will participate in Cycle 2 and beyond. The students will be asked to express their thinking with respect to particular calculus concepts by working several potential tasks. This (anonymous) work will be discussed with faculty participants to test their thinking on what it means to understand the concepts and how to elicit the appropriate understandings. The focus of this study is on the concepts, frameworks, and tasks, not student work. Student work is the medium with which we will discuss and analyze the needed calculus.

Because first-cycle interpretations are expected to be "rather barren and distorted" (Lesh & Kelly, 2000, p. 206) when compared to those that will emerge from subsequent cycles, a series of modeling cycles is usually necessary. Each cycle should challenge both the faculty members and researchers to explicitly communicate, test and assess, and reflect upon their current interpretation(s). At the end of each cycle, the investigators will have refined, reorganized,

---

mathematics occur when understanding and skill develop together and are not compartmentalized or context bound (Bruner, 1973; Tulving, 1983).

extended, or rejected their current interpretation(s). Particularly, for the researchers, data interpretation should not wait until all the data has been collected. Instead, data interpretation and analysis will take place throughout the study.

Select faculty members from an undergraduate engineering institution's mathematics and client discipline departments were recruited to participate in an iterative series of interviews centered on generating tasks that will elicit, document, and measure students' understanding of the calculus concepts the faculty participants believe to be essential to successful academic pursuits within their respective disciplines.

For the first cycle, faculty participants were grouped by course (or discipline, as necessary). These intradisciplinary groups consisted of two to four members plus a researcher. For the remaining cycles, participants will be regrouped into interdisciplinary groups to juxtapose different perspectives on calculus.

The interviews were designed around a series of concept descriptions, frameworks, and tasks developed by the researchers and/or adapted from the research of others. The intention was to provide scaffolding for the faculty to evaluate and recognize not only the necessary calculus concepts, but the ways in which the concepts need to be understood. During further interview session, research-based tasks and participant-authored tasks will serve to provide a means to elicit, document, and measure the understanding students have of these concepts.

Nineteen faculty members participated in the first interview session. Mathematics and client discipline faculty members were recruited based on their teaching experience with certain courses: single-variable differential calculus (Calculus I), integral calculus (Calculus II), and client discipline courses that listed Calculus I and/or Calculus II as a pre- or co-requisite. Of the 19 faculty participants, 6 taught math, 6 taught physics, 3 taught electrical and computer engineering, and 4 taught operations research[2] (3 of the operations research instructors were members of the Mathematics Department and the other was a member of the Management Department). Together, these instructors taught an average of 6.16 years at this particular institution, with a minimum of 2 years and a maximum of 20 years. Note: Every student at this institution, regardless of major, is required to complete an extensive core curriculum, which includes two semesters of calculus and many calculus-based science and engineering courses.

Before presenting the results and discussion for the first round of interviews, one comment must be made regarding the chosen client discipline courses. As mentioned above, the courses were selected because Calculus I and/or Calculus II was listed as a pre- or co-requisite. Surprisingly, the instructors of several of these courses would not consider them to be calculus-intensive or even calculus-based.

- One faculty member declined participation because she felt the astronautical engineering courses of interest to us did not have "sufficient material to be of use" to this study because "at no point in the semester are students actually required to perform calculus operations, [although] we talk about derivatives." For this faculty member, using calculus concepts appeared to be synonymous with computation and not with discussions or applications.
- One physics faculty participant described the delicate balance within the core physics courses: "If we go out of our way to chase all of the calculus, what we'd end up doing is

[2] Operations Research is the application of quantitative techniques to managerial decision-making.

flunking half the course; at which point we'd get a call down from the Dean. What's effectively happening is that [the incorporation of calculus is] ramping. So little by little each year, it gets a little bit tougher. So the whole class is ramping up. So in 10 years, we'll be able to say [the physics core course is] a calculus-intensive course. But certainly there's more calculus in it now than there was last year and more last year than the year before."

- The operations research faculty participants admitted that calculus is a pre-requisite "for mathematical maturity more than just the actual calculus" and because "the way [the course] is taught, you can do it without calculus."

These comments and others influence how many of the participants view calculus and the required understanding.

Results and Analysis

The first round of interviews began with a very general discussion of calculus and understanding and then progressed to the very specific. Each interview culminated in the faculty participants developing tasks to elicit the calculus understanding they discussed.

Describe Calculus.

To begin the conversation, this description of differential calculus was offered: "Linear functions are easy, non-linear functions are hard, so as much as possible, we approximate non-linear with linear." Following this, the faculty participants offered a barrage of ways to describe calculus. Calculus is …

- Combination of slopes and differences
- Study of change
- Study of infinitesimally small
- Sum of a bunch of little changes
- Tool to solve (complex) problems
- Limit as things get small
- Study of limits
- Use of limiting processes
- Subset of Linear Algebra
- Tool to explain how things work
- Math of instantaneous and continuous change
- Way to deal with a continuum world through discrete and finite concepts
- Adding up small bits of things and understanding how that sums to a bigger thing
- Language to describe what's going on in the physical world; algebra is the grammar

Despite starting from many different ideas, every group eventually described calculus as a *tool*. It is a tool to do something they needed done (e.g., explain how a physical situation works, make a prediction about some physical situation, or solve a problem).

What do you Consider to be the Fundamental Calculus Concepts?

This question was asked twice, once in general and again from the viewpoint of the participants' course(s). Like the descriptions of calculus, the list of calculus concepts began broad. The fundamental calculus concepts are …

- Function
- Infinitesimals
- Infinity
- Limit
- Derivative
- Rate of Change
- Integral
- Summation
- Accumulation
- Riemann sums
- Sequences and series
- Differential equations
- Slope
- Related rates and simultaneously changing rates

As the discussion continued, the participants narrowed this list to the "big" four: function, limit, derivative, and integral. When the discussion shifted focus to the calculus concepts used within client discipline classrooms, everyone agreed on three of these concepts: function, derivative, and integral. The disciplines differed on their view of the limit concept. For physics and operations research, the limit concept "washed out" from the list because other than "the definition of a derivative, you wouldn't necessarily worry about [limits]" and "we simple don't talk about limit in any of the courses." However, both the engineering and mathematics faculty participants highlighted the limit concept. For one math faculty participant, the limit concept is the most important of differential calculus: "the things that I want my students to take away from calculus are the behavior of functions in general and limits. I would be overjoyed if they really had a good understanding of limits. As much as anything else, I think that would be maybe the biggest thing from Calculus I."

Even though the client faculty participants listed derivative and integral as concepts they used regularly in their courses, they use different verbiage. The users of calculus deliberately use "rate of change," "summation," and "accumulation" instead of "derivative" and "integral." This choice helps them stress the physical meaning of these concepts to the students. For example, one physics faculty participant incorporates meaning and symbols by writing "sum" over every integral symbol on the board. He believes this helps the students understand that the funny "S" symbol means summation. When asked if the physics and electrical and computer engineering participants use this meaning-reinforcing verbiage in industry, they answered no. When working in industry, they use the words "derivative" and "integral" because they can safely assume their peers have the necessary understanding. But with students, they cannot make this same assumption. So, reinforcement of meaning through verbiage (and notation) is necessary in the classroom and they choose to do so with the words "rate of change," "summation," and "accumulation."

One notion that needs to be further investigated is the use of "function" by the faculty participants. Although all the participants used the term "function," we suspect the concept behind the word is equally different across groups even when the verbiage does not reflect any difference. It could be a language issue, not a content issue.

What Does it Mean to Understand a Concept? How do you Elicit This Understanding?

The assumed understanding the participants spoke of became the focus of the interview sessions, specifically a description of the calculus understanding a student needs to succeed. How each faculty participant group characterized understanding varied depending on their role as a *teacher* of calculus or as a *user* of calculus.

According to *teachers* of calculus, understanding is about the *HOW* and *WHY*. At the end of each calculus course, teachers of calculus want students to walk away with a toolbox full of tools, or procedures, the students know *how* to use, as well as *why* they should use them. Teachers of calculus want students to develop procedural fluency (i.e., to be able to carry out pre-fabricated procedures flexibly, accurately, efficiently, and appropriately). They also want students to develop procedural application (i.e., a student is able to discuss the pros and cons of a procedure, what is needed to apply one procedure versus another, and what procedure is appropriate). They believe "learning skills leads to building concepts."

How do teachers of calculus elicit the *how* and *why*? For the teachers of calculus, understanding is elicited via written tasks or oral questions. One example of such a task is given in Figure 1. This task walks a student through all the required steps to ultimately graph the given function without the aid of technology. The faculty author of this task believes very strongly that if a student can accurately produce the graph, then that student understands the derivative concept.

> Given a function: $y=2x^4+4\sin 3x,\ -3\leq x\leq 3$.
> Without using technology:
> a . Find the derivative.
> b . Find the minimum(s) and maximum(s).
> c . Find the second derivative.
> d . Find the inflection points.
> e . Where is the function increasing? Decreasing?
> f . Where is the function convex? Concave?
> g . On the following coordinate axis, graph the points from (b) and (d).
> h . Using (a), (c), (e), and (f), graph the function *y*.

***Figure 1.* Task written to elicit understanding of the derivative concept.**

The task in Figure 1 is indicative of the other tasks written by the *teacher* participants. These tasks tended to be situated in theoretical or abstract contexts, in which there are no clues given to what the function stands for, if it stands for anything. More often than not, the context is a contrived situation to get at a particular procedure.

According to *users* of calculus, understanding is about the *WHICH* and *WHY*. Within the client courses, the users of calculus want students to have more than just a toolbox full of tools, or procedures: "It's not so much that [the students] understand how to turn the crank and spit out an answer. Really mastering [calculus] relies on understanding what that integral or what that derivative actually means in the physical world." Users of calculus want students to develop *relational understanding* (Skemp, 1978/2006; i.e., recognition of the concept being dealt with and relates it to an applicable procedure). They also want students to develop *predictive power* (i.e., a student is able to step out of the mathematics and recognize that in the mathematics, there is a prediction or truth about what is happening in the physical world). Students need to assess physical situations and select the calculus tool (a pre-fabricated concept and associated procedure) that will enable them to make sensible predictions about the situation. For example, "what does the variable in this equation, that the student just constructed, mean? If it doubles, what happens to the real world? Does the student suddenly get a space ship that flies faster than the speed of light?" If so, something must be wrong with the mathematics.

How do users of calculus elicit the *which* and *why*? The users of calculus elicit understanding via written tasks. One example of such a task is given in Figure 2. Knowing that the student participants will not necessarily have any previous physics or engineering experiences, the users of calculus added some discipline-specific information to the task. Thus, the first part of this task describes an electrical field and point charge so the students are not stymied before they even start.

The electrical field of a point charge is given by □=□□□2□, where *k* is a proportionality constant, *q* is a charge, and □ is a vector from the particle to the point under consideration. Electric fields sum vectorially (superposition), so □□□□□□=□1+□2 for two particles. Using this concept, what is the electrical field at some distance *y* above the center of a horizontal bar of length *L* holding total charge *Q* evenly distributed along its length?

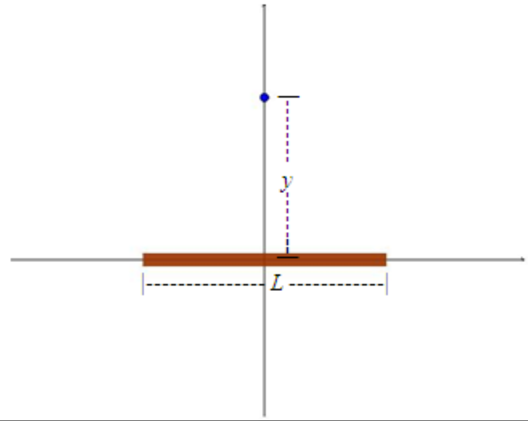


Figure 2. **Task written to elicit understanding of the accumulation concept.**

To successfully complete the task in Figure 2, students need to make a prediction about how much electric field has accumulated at the given point, which requires an integral. This task is situated in the physical world; but, still, the situation is contrived. It fits into a nice, neat little box: an *evenly* distributed charge and the point of accumulation *centered* above the bar. Even with the physical situation, the context is contrived to get at a particular concept and associated prediction.

Similarities in how both the teachers and users of calculus view understanding enable one summary to be put forth. The following presentation of how the faculty participants view understanding may be an oversimplification, it is not as cut-and-dried as it may sound. Nevertheless, the diagram is a useful characterization. According to both teachers and users of calculus, understanding is assessing the given situation and intelligently selecting an existing description (i.e., model[3]) of the expected concept and applying the associated procedures

[3] The definition of "description" used in this paper is basically synonymous with Lesh and Doerr's (2003) definition of model: "*Models* are conceptual systems (consisting of elements, relations, operations, and rules governing interactions) that are expressed using external notation systems, and that are used to construct, describe, or explain the behaviors of other system(s) – perhaps so that the other system can be manipulated or predicted intelligently" (p. 10; emphasis original). Lesh and Doerr further characterize a *mathematical* model as a model that "focuses on structural characteristics (rather than, for example, physical or musical characteristics) of the relevant systems" (p. 10).

correctly to get a reasonable answer and/or prediction. Figure 3 diagrams the process and the bulleted statements describe each corner of the 4-part diagram.

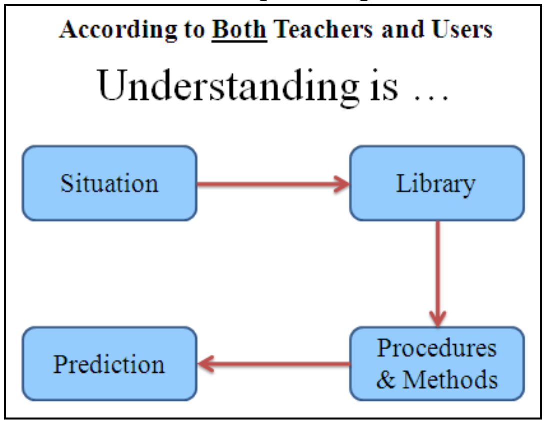


Figure 3. **4-part diagram representing both teachers' and users' views of understanding.**

- **Situation.** Whether tasks are set in a physical, theoretical, or abstract situation, the situation is massaged or formatted such that it fits into an existing description. The students' role is to recognize the expected concept.
- **Library.** Once a student has assessed the situation, he/she must intelligently select the appropriate description from those that already exist in the "library." The teacher's role is to help the students become proficient at recognizing what to "check out" from the library. The students are expected to be able to explain their check-out decision.
- **Procedures and Methods.** Procedures are linked to the choice of description(s) and the student is expected to accurately execute the appropriate procedure.
- **Predictions.** Based on the procedure's result, the student is to make a sensible prediction about the given situation. This includes asking if it makes sense. For example, did the units work out? Would changing a variable yield the expected result?

The teachers and users of calculus are not giving a situation to students and asking them to use calculus to describe and explain that situation. They are simply using contrived situations to help students become better at choosing a description that is already developed and/or constructed for them and then using the linked procedures accurately. The students are not really looking into the situation and using calculus to describe it; instead they are asked to select a description that works.

Discussion and Conclusion

Throughout the first round of interviews, the following thoughts and comments emerged as noteworthy. "Pattern matching," as some participants called it, allows some students to really master calculus, to really understand what the integral or derivative means in the physical world. But for a good majority of students, pattern-matching does not work. Something is missing. One participant asked the researcher if this study was going to change the way calculus was being taught because, from his viewpoint, many of the students were not getting what they needed to succeed outside the calculus classroom. According to the users of calculus, an A in calculus does not mean the student understood and could apply calculus to the physical world. But, according to the teachers of calculus, earning an A means the student really understands deeply. If the purpose of Calculus I and Calculus II is to prepare students for client discipline courses, something needs to change.

Analyzing the interviews and much research, the researchers were led to believe that if someone is using calculus very sensibly (i.e., with a great deal of understanding), he/she can assess a situation and use calculus concepts to describe it. The calculus concepts are linking to procedures, which are then used correctly to make predictions. Unlike before, the process does not stop here. When the individual gets the result and/or prediction, he/she will know what assumptions were made and what things might put some spin on what can be done with the result. This spin may imply some margin of error, degree of uncertainty, or difference made by making one assumption versus another. Those are things that someone who just checks-out something from the library is never going to understand because they never had to. Figure 4 diagrams this hypothesized process, with the differences emphasized with dashed lines. The bulleted statements describe each piece of the 4-part diagram.

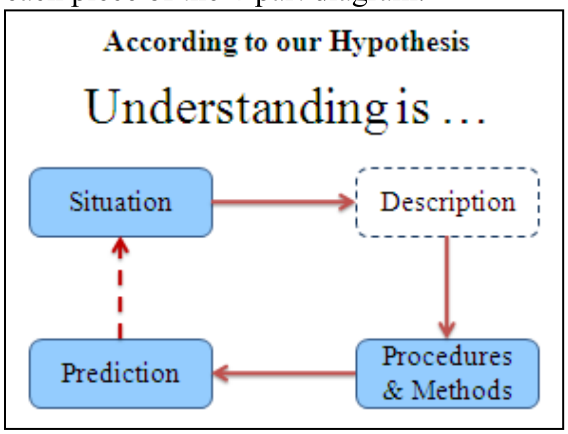


Figure 4. **4-part diagram representing our hypothesis view of understanding.**

- **Situation.** Realistic situations are given to students. These situations usually do not quite fit into an existing description.
- **Description.** Students look at a situation, in the raw, and use calculus concepts to compose a description, not necessarily one that matches something in the library.
- **Procedures and Methods.** The concepts used to describe the situation are linked to procedures, which the students are expected to use correctly.
- **Predictions.** Based on the result of the procedure, the students are to make a sensible prediction about the given situation.
- **Relate to Situation** (dashed arrow)**.** The prediction came from making assumptions which spin what can be done with the result. Implies some margin of error, degree of uncertainty, difference made by making one assumption versus another.

Thinking of understanding in this fashion can be likened to cooking. One level of cook will go to the cupboard and assess the ingredient situation. Say, he/she finds tomatoes, mozzarella, and oregano. To this cook, this means an Italian meal and so he/she goes to the library, finds an Italian cookbook and selects a recipe that has only the ingredients in the cupboard. This cook follows the recipe, step-by-step, and produces a meal. It may be an excellent meal, but the cook cannot modify it if needed. If the recipe calls for tomatoes and they are not in season, this cook is forced to find a different recipe because he/she does not know how to modify the recipe. But another level of cook, one with a great deal of understanding, will assess the ingredient situation and compose a recipe instead of choosing one. This person is always going to be a better cook because he/she can take whatever is in the refrigerator, what is fresh and in season, and create an excellent meal. He/she does not have to use tomatoes even when they are not in season.

This second level of cook has the type of flexible, durable, and applicable understanding we want from calculus students. We want a student that can use calculus to describe a realistic situation, one that might not fit any existing descriptions, textbook examples or library entries. When a student arrives at a result, they will know if and how it applies to the given situation and whether it applies to any other situation. This occurs because they know what assumptions they made, what error or uncertainty might be involved. Because the calculus description was created by the individual, he/she owns it. This personal ownership enables him/her to not only build and deepen their understanding, but to use calculus in novel and unfamiliar situations of any kind. This type of understanding is flexible, durable, and transferrable. One faculty participant summed it up this way: "Ultimately we want, not get [students] to be able to do problems similar to what they've seen before, we want them to take what they've known and do something new."

As might be expected, the tasks that elicit this level of understanding look quite different than those presented earlier. Two example tasks are below. The Toy Train Problem in Figure 5 asks students to design a toy train display. They are asked to make the mountain as tall as possible, use the least amount of track, and not exceed a 10% grade on the track. The central calculus concepts students need to use in this problem are rate of change (i.e., derivative) and accumulation (i.e., integral).

The Toy Train Problem

At Christmas time last year, the Indianapolis Children's Museum built a number of large displays using toy trains, gingerbread houses, plaster mountains, and other small toys. Each display covered the top of a large table that was at least 2 meters wide and 3 meters long; and, each display included three or four different sized mountains that were made using paper and plaster. The goals were: (a) to make the mountains as tall as possible, (b) to minimize the amount of train track that is used, and (c) to make sure that the grade of the track never exceeds 10% going up or going down. … But, last year, one huge problem occurred! The toy trains were only powerful enough to go up or down a slope of 10% or less. And, last year, the slopes of most of the tracks were too steep for the trains to run.

The topographic map on the back of this page shows the heights, locations, and shapes of the mountains for one display that was used last year. The table for this particular display was 2 meters by 3 meters; and, the contour lines show the height of the mountains. For example, at the edges of the darkest regions on the map, the height of the mountains is 200 centimeters; at the edges of the next darkest regions on the map, the height of the mountains is 175 centimeters; and so on. … The map also shows the location of the toy village and Santa's Workshop. But, the map does not show the path of the train track that was used. This is because the path of the train track that was used last year was too steep for the trains to run.

<u>Your Task</u>: Write a letter to the people who will be designing the toy train displays this year. Describe how to decide how tall the mountains can be on each table; and also describe if the shapes or locations of some of the mountains need to be changed. To begin, re label the contour lines on the topographic map to show how tall the mountains cane for the display that

is shown on the back of this page. Then, show a path for the train tracks so that the grade will never exceeds 10% going up or down the mountain. Finally, draw diagrams and describe how similar decisions can be made for other tables that have different dimensions. … The display designers need your help as fast as possible. You only have 60 minutes to write your letter and submit your plans.

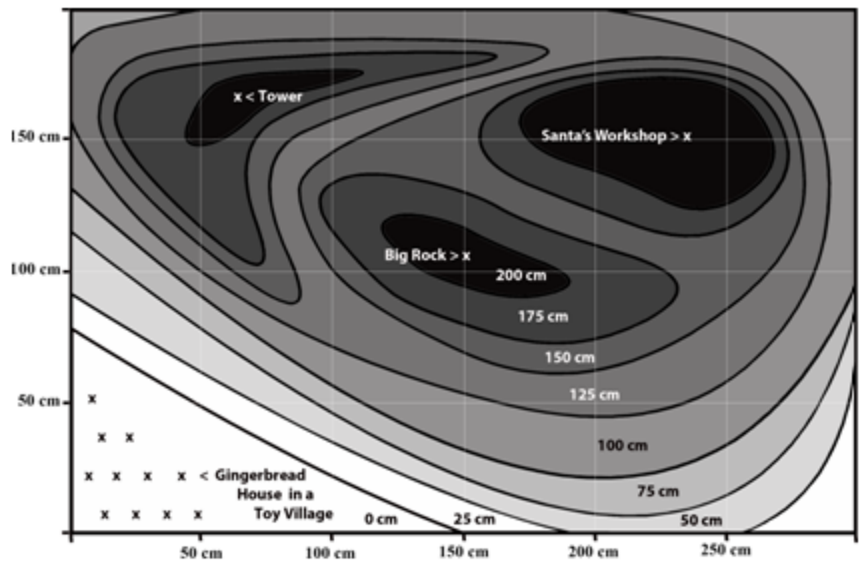


Figure 5. The Toy Train Problem.

The Theater Problem in Figure 6 asks students to create a recommendation for when patrons should arrive at a movie theater to minimize their wait time. Using the given rates and numbers of attendees, the students are asked to minimize the wait time experienced by patrons, while maintaining a 10-minute buffer for "settling in to the movie." The central calculus concepts students need to use in this problem are rate of change (i.e., derivative), accumulation and area (i.e., integral), and the relationship between derivatives and integrals.

The Theater Problem

During last year's summer blockbuster season, the Cinema 12 Theater had horrendous lines at the ticket counter. On several occasions, patrons missed the beginning of a movie or did not have time to buy popcorn and drinks for the show. This lost a ton of revenue for the theater. At other times, patrons waited outside a theater while the cleaning crew finished cleaning up after the previous showing. With the blockbuster season effectively kicking off in a couple of weeks, the theater owner would like to post recommendations on the theater's website regarding how long before a particular showing patrons should arrive. The ideal arrival time should give the patrons time to purchase tickets and leave a 10-minute window for patrons to "settle in" for the movie (e.g., buy popcorn and drinks). The owner is asking for your help in creating these arrival-time recommendations.

The theater owner reviewed last year's records. He found that all the shows between 6 and 9 pm sell out on opening weekend. After that, each showing attracts 75% of each theater's capacity. Each theater has seating for 100 people. Matinee shows (those before 6 pm) attract 40% and late night shows (those after 9 pm) attract 65% of each theater's capacity. To sell tickets, the owner has 2 attendants working at all times, with 2 more standing by, ready to work, depending on the length of the patron line. Each attendant takes no more than 30 seconds to process one patron's request. While the owner wants to minimize the patrons' wait time, he also needs to minimize the number of attendants working.

Your Task: Based on this information, and an example of show times listed below, write a letter to the theater's owner giving your proposed arrival-time recommendations and the number of attendants the owner can expect to work throughout the day. In your letter, describe how you arrived at each recommendation and expectation. Be sure to give details and a clear explanation so the owner can decide whether your recommendations are the ones he should post.

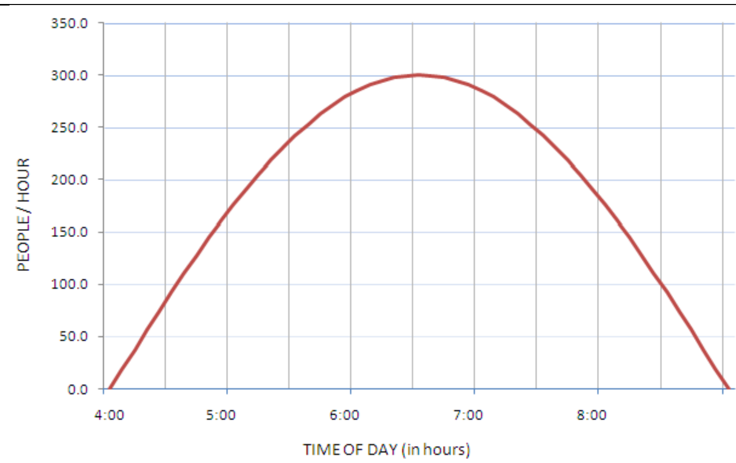
Show times for the Cinemas 12 Theater.

Movie #1	11:10 am	1:50 pm	4:30 pm	7:30 pm	10:25 pm
Movie #2	11:20 am	2:10 pm			
Movie #3	10:45 am	1:30 pm			
Movie #4	10:40 am				
Movie #5	11:50 am	2:30 pm	4:45 pm	7:00 pm	9:30 pm
Movie #6	4:15 pm	7:45 pm	10:30 pm		
Movie #7	10:30 am	1:15 pm	4:00 pm	7:15 pm	10:00 pm
Movie #8	11:30 am 7:10 pm	1:20 pm 8:00 pm	2:20 pm 10:15 pm	4:10 pm 10:45 pm	5:10 pm
Movie #9	11:00 am	1:45 pm	4:20 pm	6:50 pm	9:50 pm
Movie #10	4:50 pm	7:40 pm	10:20 pm		
Movie #11	9:40 pm				
Movie #12	11:40 am	2:15 pm	4:40 pm	7:25 pm	9:45 pm
Movie #13	10:50 am	1:40 pm			
Movie #14	10:20 am	1:10 pm	4:05 pm	6:45 pm	
Movie	11:15 am	2:00 pm	4:35 pm	5:00 pm	7:20 pm

		7:50 pm	10:10 pm	10:40 pm		
The times above are example show times (start times for each movie).						

Figure 6. The Theater Problem.

The Theater Problem is a modelling version of a task written by one of the faculty participant groups. The original task (see Figure 7) did not involve varying rates because one of the faculty-author's understanding of cuing theory, which she was currently teaching, did not allow for anything but constant rates. Her faculty partner, an electrical engineer by training and a mathematician through teaching, commented that "it is such a shame to have her say this because having a varying rate is why we need calculus." The levels of calculus understanding demonstrated within this group are what we are trying to elicit, document, and measure with the tasks.



The chart above shows the arrival rate of people coming to a theater box office to buy tickets. The box office has the capacity to sell tickets at a max rate of 200 people/hr.

a . At what time did the line begin to form?
b . At what time was the line the longest?
c . How long is the line when it is at its longest?
d . Is there still a line at 9:00?
e . If you arrive at 6:00, how long will you wait?

Figure 7. **The theater problem, as written by one faculty participant group.**

Both The Toy Train Problem and The Theater Problem stem from realistic, raw situations that require students to use calculus concepts to describe what is happening and to arrive at a prediction that can be extended to other situations. Note that neither problem specifically requires students to calculate a derivative or integral or even model the situation with a function. This distinction might categorize these tasks, in the minds of faculty members such as the astronautical engineering instructor that declined to participate, as non-calculus tasks. For them, a task that does not require calculus computations is not a meaningful task. These instructors are comfortable with traditional textbook word problems which emphasize computational skills, as evidenced by the tasks written by the participants during the first interview session to elicit calculus understanding.

For those familiar with the Models and Modeling Perspective (Lesh & Doerr, 2003; Lesh, Hamilton, & Kaput, 2007; Lesh, Hoover, & Kelly, 1993; Lesh & Lamon, 1992; Lesh, Landau, & Hamilton, 1983), the tasks in Figures 6 and 7 are very similar to problem-solving activities often referred to as *model-eliciting activities* (*MEAs*). MEAs require students to "go beyond short answers to narrowly specified questions – which involve sharable, manipulatable, modifiable and reusable conceptual tools (e.g., models) for constructing, describing, explaining, manipulating, predicting, or controlling mathematically significant systems" (Lesh & Doerr, 2003, p. 3). MEAs differ from traditional word problems in several ways (see Figure 8). First, in MEAs, students create mathematical descriptions (i.e., models) of meaningful situations (i.e., real world, raw situations). In traditional word problems, students make meaning of symbolically described situations. This difference is the difference between mathematizing and decoding (Lesh & Doerr, 2003). By mathematizing, Lesh and Doerr mean "quantifying, dimensionalizing, coordinatizing, categorizing, algebratizing, and systematizing relevant objects, relationships, actions, patterns, and regularities" (p. 5). A second difference is what assumptions are made.

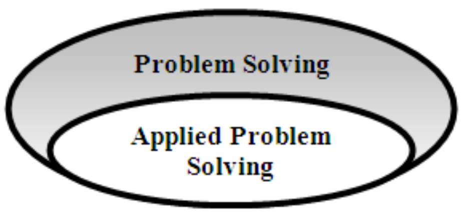
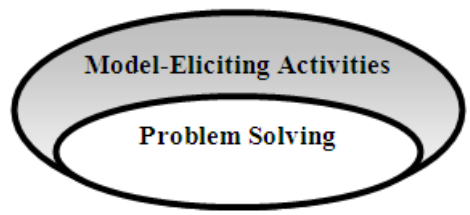
Traditional Views Applied problem solving is treated as a special case of traditional problem solving.	**Modeling Perspective** Traditional problem solving is treated as a special case of model-eliciting activities.
	
Learning to solve "real life" problems is assumed to involve three steps: 1 . First, learn the prerequisite ideas and skills in decontextualized situations. 2 . Learn general content independent problem-solving processes & heuristics. 3 . Finally (if time permits), learn to use the preceding ideas, skills, and heuristics I messy "real life" situations where additional information also is required.	Solving meaningful problems is assumed to be easier than solving those where meaningful interpretation (by paraphrasing, drawing diagrams, and so on) must occur before sensible solution steps can be considered. Understanding is not through of as being as all or nothing situation. Ideas develop; and, for the constructs, processes, and abilities that are needed to solve "real life" problems, most are at intermediate stages of development.

***Figure 8.* Applied problem solving ≠ model-eliciting activities (Lesh & Doerr, 2003, p. 4).**

A secondary purpose of this study is to provide calculus instructors with tasks that can used to measure student understanding of calculus. Once the instructor has this measurement, he/she can adjust his/her instruction to better move the students from where they are to where they need to be. So, in a sense, this study is trying to change the way calculus is currently taught.

This measurement can happen either before or after instruction. If used before instruction, these tasks encourage students to develop their own understandings; and if used after, the student can apply what they have already been taught (Lesh, Yoon, & Zawojewski, 2007). As Yoon, Dreyfus, and Thomas (2010) suggest, if tasks like those described here are implemented after instruction, they not only measure and document a student's understanding, but allow the student to deepen his/her understanding of the calculus concept(s) while affording an opportunity to apply his/her understanding in a realistic situation. Research has shown that for MEAs, students often invent (or significantly extend, refine, or revise) constructs that are more powerful than anybody has dared to try to teach to them using traditional methods (Lesh, et al., 1993).

As stated before, calculus is an important tool for building mathematical models of the world around us and is thus used in a variety of disciplines, such as physics and engineering. These disciplines rely on calculus courses to provide the mathematical foundation needed for success in their discipline courses. Therefore, at its conclusion, this study hopes to offer a collective vision to focus the content of beginning calculus courses on the meeting the needs of client disciplines. However, in the end, it is the mathematicians that have the responsibility to create courses and curricula that embrace the spirit of this vision while maintaining the intellectual integrity of mathematics. By explicitly knowing what and how students should be prepared for client courses, teachers and curriculum developers of both calculus and client disciplines can work together to prepare students for academic success in any discipline.

REFERENCES

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141-178.

Bruner, J. S. (1973). *Beyond the information given: Studies in the psychology of knowing* (1st ed.). New York, NY: Norton.

Collins, A. M. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22). Berlin; New York: Springer-Verlag.

Douglas, R. G. (1986). The Importance of Calculus in Core Mathematics. *Journal of College Science Teaching, 15*, 250-251.

Ferrini-Mundy, J., & Gücler, B. (2009). Discipline-based efforts to enhance undergraduate STEM education. *New Directions for Teaching and Learning, 117*, 55-67.

Ganter, S. L., & Barker, W. (2004). A collective vision: Voices of the partner disciplines. In MAA (Ed.), The Curriculum Foundations Project (pp. 1-13). Washington, D.C.: Mathematical Association of America's Committee on the Undergraduate Program in Mathematics.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., . . . Human, P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.

Hughes Hallett, D. (2000). *Calculus at the start of the new millennium.* Paper presented at the Proceedings of the International Conference of Technology in Mathematics Education, Beirut, Lebanon.

Lesh, R. A. (2002). Research design in mathematics education: Focusing on design experiments. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 27-49). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lesh, R. A., & Doerr, H. M. (2003). Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. A. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp. 3-33). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lesh, R. A., Hamilton, E., & Kaput, J. J. (Eds.). (2007). *Foundations for the future in mathematics education*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lesh, R. A., Hoover, M., & Kelly, A. E. (1993). Equity, assessment, and thinking mathematically: Principles for the design of model-eliciting activities. In I. Wirszup & R. Streit (Eds.), *Developments in school mathematics education around the world: Volume 3* (pp. 104-130). Reston, VA: National Council of Teachers of Mathematics.

Lesh, R. A., & Kelly, A. E. (2000). Multi-tiered teaching experiments. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 197-230). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lesh, R. A., & Lamon, S. J. (1992). *Assessment of authentic performance in school mathematics*. Washington, D.C: American Association for the Advancement of Science.

Lesh, R. A., Landau, M., & Hamilton, E. (1983). Conceptual models in applied mathematical problem solving. In R. A. Lesh (Ed.), *The acquisition of mathematical concepts and processes* (pp. 264-344). New York, NY: Academic Press.
Lesh, R. A., Yoon, C., & Zawojewski, J. (2007). John Dewey revisited -- Making mathematics practical versus making practice mathematical. In R. A. Lesh, E. Hamilton & J. J. Kaput (Eds.), *Foundations for the future in mathematics education* (pp. 315-348). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Lovell, M. C. (2004). *Economics with calculus*. Hackensack, NJ: World Scientific Pub.
Skemp, R. R. (1978/2006). Relational understanding and instrumental understanding. *Mathematics Teaching in the Middle School, 12*(2), 88-95.
Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.
Yoon, C., Dreyfus, T., & Thomas, M. O. J. (2010). How high is the tramping track? Mathematising and applying in a calculus model-eliciting activity. *Mathematics Education Research Journal, 22*(1), 141-157.

USING TOULMIN ANALYSIS TO LINK AN INSTRUCTOR'S PROOF-PRESENTATION AND STUDENT'S SUBSEQUENT PROOF-WRITING PRACTICES.

Timothy Fukawa-Connelly
The University of New Hampshire
tim.fc@unh.edu

This paper provides a method for analyzing undergraduate teaching of proof-based courses based on Toulmin's model of argumentation. The paper presents a case-study of one instructor's presentation of proofs and shows that she was inconsistent in the amount of detail that she included in her written proofs. The paper then describes how that analysis can be used as a predictor of subsequent student proof-writing performance. Second, student work is analyzed using Toulmin's (1969) model for argumentation. The data shows that the students had adopted the modeled proof-writing behavior of the instructor in terms of the type and quantity of elements of argumentation. The method of analysis was developed via research in a lecture-based abstract algebra class, but has applications to any lecture-based proof-intensive course. This method provides one way to link classroom teaching activities to student performance that forces instructors to assume more responsibility for their students' demonstrated end-of-course performance.

Keywords: Toulmin's model; abstract algebra; proof-presentation; undergraduate mathematics teaching

Introduction and Research Questions

There are many critiques of lecture-based mathematics instruction, especially in the context of proof-based courses, which argue that lecture-based teaching is intimidating and that it misleads students about the nature of mathematics (Thurston, 1994; Cuoco, 2001). Others contend that it hides much of the process used in mathematical thinking and makes it difficult for students to learn or develop an appreciation for the discipline (Dreyfus, 1991). Some maintain that it ignores the important role that mathematicians ascribe to ideas such as elegance, intuition, and cooperation (Burton, 1998; Dreyfus, 1999; Fischbein, 1987). It has also been claimed lecture-based instruction is that it is not an effective way to promote student learning of the mathematics content (Leron & Dubinsky, 1995). In fact, there is evidence that undergraduate mathematics teaching is one of most important reasons that students change majors away from STEM fields (Seymour & Hewitt, 2004).

While each of these assertions may be true, "very little empirical research has yet described and analyzed the practices of teachers of mathematics" (Speer et al., 2010, p. 99) at the undergraduate level despite repeated suggestions for the study of university mathematics teaching (Harel & Sowder, 2007; Harel & Fuller, 2009. For example, Mejia-Ramos and Inglis (2009) also conducted a literature search and found only 102 papers that were educational research studying undergraduate students' experiences writing, reading and understanding proof. Of those 102 papers, they found no papers describing how students understand instructor's presentation of proof. That is, "researchers' questions, methods, and analyses have not generally targeted what teachers say, do, and think about collegiate classrooms in an extensive or detailed way" (Speer et al., 2010, p. 105). In particular this paper will:

1) Use Toulmin's (1969) model of argumentation to analyze the proof-presentation of an instructor of undergraduate abstract algebra course.

2) Propose that students in the class will adopt argumentation methods that are modeled by their instructor.

3) Analyze student work to determine whether the Toulmin analysis of teaching does, in fact, predict students' proof-writing behavior.

Studying Teaching

1.1 Studying undergraduate mathematics instruction

There has been little effort to study what actually is happening in college mathematics classrooms, especially lecture-based classrooms (Speer, Smith, & Horvath, 2010). This gap remains despite numerous suggestions for such study (Harel & Sowder, 2007; Harel & Fuller, 2009).

As Mejia-Ramos and Inglis (2009) documented there are few studies that have analyzed the teaching of undergraduate proof-based courses. Weber (2004) described one instructor's teaching of real analysis and classified his presentation of proofs into three categories. Alcock (2009) interviewed five mathematicians about their teaching of an introduction to proofs class and described the types of thinking that they attempt to inculcate in their students along with some suggestions of how they might attempt it. Weber (2010) has similarly interviewed mathematicians about the reasons that they might present proofs and how that relates to their teaching of an introduction to proofs class. Yet, "none of the articles discussed tasks directly focussed on the presentation of an argument to demonstrate students' understanding of it" (Mejia-Ramos & Inglis, 2009, p. 91). In particular, Mejia-Ramos and Inglis argue that, "we, mathematics educators, know very little about students' behaviour in some of the main types of activities involved in the assessment of their proving skills" (2009, p. 93) and argue that research should focus on analyzing the presentation of proofs and how students then understand them.

The research that has been done on advanced mathematics courses has shown that lecture-based instructors consciously model important mathematical behaviors and ways of thinking as part of their teaching (Weber, 2004; Alcock, 2009; Fukawa-Connelly, 2010). I posit that students appropriate some of the modeled behavior, but not always those aspects that faculty believe are most important. This argument is supported by research on transfer-in-pieces (Wagner, 2006). This research claims that novices may not perceive the same aspects of a demonstration or problem as important that an expert would, and instead attend to other, less important aspects. Wagner (2006) demonstrated that students subsequently transferred aspects of a structure from one problem to another that were at times inappropriate.

In a proof-based course like abstract algebra, the instructors model proof-writing strategies and the types of arguments that they expect from their students. We can analyze these instructors' proof-presentation to understand what they expect of students and look to see if students do, in fact, display similar proof-writing habits. It is necessary to keep in mind that the students may have appropriated aspects that were different than what the instructor intended.

After analyzing classroom teaching, this paper turns to student work. The paper provides a preliminary means of linking the analyses of classroom instruction with student learning. To analyze the way the teacher modeled proof-writing, this paper draws upon Toulmin's model of argumentation (1969) as an analytical tool. In particular, this paper focuses on the relationship between the instructor's modeled presentation of data, warrants, backing, and qualifiers and examines how that presentation is reflected in student work. I show that the students' written proofs are similar to those presented by the teacher. Admittedly, this is an imperfect measure, as the students also receive written feedback on homework and examinations, but it is reasonable to assume that the behavior careful instructors model in the classroom is consistent with their written commentary.

This analysis also suggests that analyzing the instructor's classroom presentation of proof is a valid way to explain student work. It reinforces the often-heard refrain, that "students learn what you teach them." Finally, while this analysis is presented in the context of an algebra course I argue that it is applicable to any proof-based mathematics course and especially useful in analyzing links between lecture-based teaching and student proof-writing.

1.2 What do I mean by lecture?

Let me briefly specify what I mean by traditional, lecture-based courses. Rasmussen and Marrongelle (2006) described a scale of teaching that ranges along a continuum from "pure telling" to "pure investigation." Similarly, McClain and Cobb (2001) characterized the spectrum of teaching as running from "non-interventionist" to "total responsibility." The instructor for the class described in this study used a style closer to "pure telling" than "pure investigation," and had almost total responsibility for daily classroom activities.

Toulmin's Model-Global and Local Arguments

Significant research has been done in the past decade drawing upon Toulmin's (1969) model of argumentation to analyze student proof and justification in mathematics class (Krummheuer, 1995; Krummheuer, 2007; Weber, Maher, Powell, & Lee, 2008; Inglis, Mejia-Ramos, & Simpson, 2007; Yakel, 2001). In particular, these researchers have used Toulmin's model to analyze and explain first-grade classroom discussion of number decompositions (Krummheuer, 2007), sixth-grade discussion of warrants in statistical arguments (Weber, et al., 2008) and a high school geometry class (Knipping, 2008). Krummheuer's (2007) goal was to describe the structure of the argument produced in classroom interaction and the different ways students were involved in the production of that argument. Knipping (2008) extended Krummheuer's work to encompass more sophisticated mathematical work. Yet, even then, Knipping made no effort to link possible student learning to classroom activity. Inglis, Mejia-Ramos, and Simpson's (2007) work suggested that analysis of mathematicians' proof production requires drawing upon Toulmin's full model as opposed to the reduced version more commonly used in mathematics education research. Toulmin's scheme has six basic types of statement, each of which play different roles in arguments. These are:

- The data (D): the foundation upon which the argument is based
- The conclusion (C): that which is being argued
- The warrant (W): justifies the relationship between the data and the conclusion
- The backing (B): supports the warrant by suggesting why it is valid, or, put another way, explains the permissibility of the warrant.
- The modal qualifier (Q): expresses a degree of confidence of the conclusion
- The rebuttal (R): states conditions under which the conclusion would not hold.

Traditional instructors typically model the mathematical behavior that they expect from their students (Fukawa-Connelly, 2010). In this case, the behavior is limited to proof-writing skills, including the level of detail that the instructor believes to be appropriate. Taken seriously, the idea that instructors model the level of detail they want students to include in arguments suggests that what the teacher does in class serves as a model for the level of detail that the students are to include in their written proofs. I propose that this includes whether, and when, proofs include explicit statements of warrants, backing, qualifiers and rebuttals. This paper will make explicit one connection between instruction and student work, but it does not make any effort to comment on the appropriateness of the instruction or the instructional goals. It should not be read, in any way, as a critique of the instructional methods. It is meant to provide a means to analyze teaching and student work and give an explanatory link between them.

Data and Methods

3.1 The teacher and institution

When the study began, Dr. Tripp (a pseudonym) was an assistant professor working towards tenure at a mid-sized doctoral granting institution in the Midwest. Dr. Tripp holds a doctorate; her research specialty is in algebra. She had taught abstract algebra a number of times prior to the study. In the context of this study, Dr. Tripp's most important characteristic is her self-description as a traditional teacher of abstract algebra. For her, this meant that lecturing in front of the class would be the predominant method of instruction, that "proofs form the backbone of this course," and that her organization and presentation of material would closely follow that of the text, *Abstract algebra: An introduction* (Hungerford, 1997). In other words, on the continuum of teaching from pure telling to pure discovery, Dr. Tripp claimed that she would be closer to pure telling.

3.2 The class

The class met four times per week for 50-minute sessions in the spring semester. Students had frequent homework assignments from the text, two in-class exams, and one final exam. The course featured a significant amount of ring theory, starting with the definition of a ring, moving onto quotient rings, and culminating with the construction of roots of irreducible polynomials. The class transitioned to the study of group theory with approximately one week remaining in the semester.

3.3 The students and data collected from students

Typically, the students were juniors who had completed a two-semester calculus sequence. All studentswere required to complete an introductory course on mathematical proof as prerequisite to abstract algebra; thirteen such students were enrolled in the course simultaneously in the section studied, including one sophomore and one senior. Twelve students agreed to participate passively in the study by allowing recording, transcription and analysis of their classroom activities. Five students actively participated by completing three written assessments. The first was a survey that asked them to summarize their mathematical preparation; this was followed by two content surveys, the first of which will be discussed in this study.

3.3.1 The students of the class.

There were five students from the class who agreed to actively participate in all aspects of the study (all identified via pseudonyms in work below). All of the students who agreed to participate had an overall GPA ranging from 3.6 to 3.9, which was representative of the students enrolled in abstract algebra that semester, according to the faculty. Four were junior mathematics majors, one was a mathematics education major and two also had a second major. All had earned either an A or B in their introduction to proofs class. However, the faculty did not believe it had been an effective course. Thus, I argue that it did not have a significant influence on the students' proof-writing. Finally, one active participant was a sophomore who was also doing independent research in graph theory with another professor in the department.

She had won the departmental freshman/sophomore prize in mathematics as both a freshman and a sophomore.

3.3.2 Content assessment instrument.

The content assessment instrument analyzed in this paper was given mid-semester. It presented a single task that assessed whether the participating students were able to determine if a proposed set with associated operations was a ring, which is a typical exercise in an introductory algebra course. The task, as given to the students, is shown in Figure 1.

Let $(\mathbf{Z}, +, \times)$ be the ordinary ring of integers: $\mathbf{Z} = \{\ldots -2,\ -1,\ 0,\ 1,\ 2,\ \ldots\}$, $n+m$ is the sum of n and m, $n \times m$ is the product of n and m. If n and m are in $\mathbf{Z}$ then $n \geq m$ if and only if $n-m$ is either positive or 0.

The algebraic system $R = (\mathbf{Z}, \oplus, \otimes)$ consists of the set $\mathbf{Z}$ with the two binary operations $\oplus$ and $\otimes$. The operation $\oplus$ is ordinary addition: $n \oplus m = n + m$. The operation $\otimes$ is defined by: $n \otimes m = \begin{cases} n & \text{if} \quad n \geq m \\ m & \text{if} \quad m \geq n \end{cases}$. Otherwise put, $n \otimes m$ is the maximum of n and m.

You are to determine if R is a ring. If it is a ring, show that it satisfies all of the defining properties of a ring. If it is not a ring, identify a needed property that R fails to satisfy, and demonstrate that this property is not satisfied by R.

Figure 1. Task given to students used in analysis

3.4 Methodology for analysis

I observed 18 of the class meetings, taking detailed field notes, and made video recordings of 15 of those classes. I transcribed all dialogue as well as all text on the board. The video camera primarily recorded the instructor, as she was the principle focus of all class activities. I reviewed all classroom video recordings and made a log of all episodes that included proof-writing or presentation. In all of the examples described below, pseudonyms are used to preserve the anonymity of the participants. These pseudonyms are consistent throughout the paper; all work and comments by Jeff, for example, are from the same person. In the transcript of classroom data, most student utterances are simply denoted via **S**, without regard for the student speaking, while all of Dr. Tripp's begin **Dr. Tripp.**

Criteria for proof-production or presentation was straightforward; an incident was logged as such when any member of the class community wrote or showed a formal mathematical proof that drew on symbolic notation and logical reasoning. This excluded any informal justification that was offered if it was not at that moment connected to formal proof. Work with particular examples of structures was categorized as proof as long as algebraic techniques were being used. I was inclined to count written work alone as modeling of proof-writing, simply because it was more likely to be copied into students' notes, but have also included some analysis of the spoken modeling of argumentation.

After identifying proof-writing episodes in class, I noted the following pair of proofs demonstrating that a function, $\overline{f}$, is a homomorphism of rings. The first proof was written by Dr. Tripp and the second by a student, Jeff, immediately after Dr. Tripp wrote her proof.

Dr. Tripp's work	*Jeff's work*
$\overline{f}((r+\ker f)+(t+\ker f))$	$\overline{f}((r+\ker f)(t+\ker f))$
$=\overline{f}((r+t)+\ker f)$	$=\overline{f}((rt)+\ker f)$
$=f(r+t)=f(r)+f(t)$	$=f(rt)=f(r)f(t)$
$=\overline{f}(r+\ker f)+\overline{f}(t+\ker f)$	$=\overline{f}(r+\ker f)\overline{f}(t+\ker f)$
$\therefore \overline{f}\ \ preserves\ \ addition$	$\therefore \overline{f}\ \ preserves\ \ multiplication$

That Jeff had perfectly reproduced Dr. Tripp's work was immediately apparent; it seemed obvious that he was using Dr. Tripp's proof as a model for his own. That is, I believed that Jeff had appropriated the manner of writing proof that Dr. Tripp had modeled in her work.

Similarly, Dr. Tripp wrote out a proof demonstrating that a function preserves addition, then called Nathan to the board. He wrote a near-perfect copy:

Dr. Tripp's work	*Nathan's work*
$f(a,b)+f(c,d)$	$f(ac,bd)$
$=b+d$	$=bd$
$=f((a+c,b+d))$	$=f(a,b)f(c,d)$

It is important to note that these are the only two cases where Dr. Tripp wrote an argument and then had a student write an argument about a similar statement. There were other cases where students gave proofs at the board or using an overhead transparency, but none were similar to Dr. Tripp's work.

Thus, to pursue the question of whether students had appropriated proof-writing behavior that Dr. Tripp modeled required analysis of the proofs she wrote during class and of the students' work outside of class. She, collectively with the students, wrote one complete proof of the fact that a structure is a ring, verifying each of the required properties and then showing that the ideal property holds. She wrote one proof showing that a structure is a subgroup, three proofs demonstrating that a function is a homomorphism, and proofs that the kernel is a given set, that a function is onto and that a function is one-to-one. Additionally, I observed her or a student present five property-verification arguments: that a given function is well-defined, that the quotient structure R/I is associative under addition, that multiplication distributes over addition when R is a ring and I is an ideal of the ring, that all non-zero elements of a given ring have multiplicative inverses, and, finally, that a given function is onto a specified domain. Finally, I observed 6 presentations by students that were prepared ahead of time, which essentially were proofs copied from the text and thus limited in their usefulness to the present analysis.

In sum, I observed five arguments that required a number of independent properties had to be demonstrated in order to show that a structure is as desired. I also observed 23 individual property verification arguments across these reserve-structure proofs and the other arguments described above. Because I only have two examples of students writing proofs involving homomorphism property verifications, both shown above, I decided to concentrate my analysis of Dr. Tripp's teaching on ring and group property verification arguments.

I analyzed all arguments using Toulmin's (1969) model, noting both when an aspect of argumentation was written and when one was spoken aloud. In this paper I Dr. Tripp's demonstration of all of the properties required for a sub-ring proof.

All of Dr. Tripp's written and spoken statements used in proof-writing were classified as one of the following: data, warrant, backing, qualifier or conclusion. The students' written work was similarly classified. A statement was classified as data if it directly followed "if" or "let" or was stated as an obvious fact without support (such as $f(0)=0$ when discussing a homomorphism, because this is an elementary fact about morphisms). A statement was classified as a conclusion when it followed "then" or when it was presented as a deduction from a piece of data, such as "R fulfills X property." A statement was classified as a warrant when it linked the data and conclusion in a way that explained how the data supported the conclusion and drew upon previously demonstrated facts, or upon facts that had been stated as part of the hypotheses (for instance, information about elements included in sets). A statement was classified as backing if it explained the applicability of a warrant such as the claim "We have previously proved…" Finally, a statement was classified as a qualifier if it limited the generalizability of the statement, provided a possible exception to the statement, or asserted facts about logic that served a similar purpose. In complex proofs some statements are first introduced as conclusions and then reused as data for a subsequent argument.

Dr. Tripp described herself as a traditional, lecture-based teacher, but it is important to note that in her classroom, the students were likely to participate in proof-writing or presentation activities. In 29 observed proofs, the students presented seven by themselves. Dr. Tripp used participatory proof-writing schemes 21 times, engaging students in a question-and-answer dialogue as she was working. Only once did she deliver the caricatured "lecture" where students sit quietly and take notes without questions or comments while the professor presents the proof. Thus, her classroom was filled with dialogue, and, in order to analyze her presentations of proofs, it is necessary to recognize that this dialogue plays an important role in her pedagogy. Thus, I created a summary table that captured how frequently in these arguments Dr. Tripp spoke, wrote, or both spoke and wrote or asked a student to articulate a particular aspect of an argument. I employed a similar methodology to analyze the students' written work.

Analysis

4.1 Analysis of teaching property verification arguments

Property verification arguments were more readily accessible for analysis as Dr. Tripp always wrote out substantial portions of them. Similarly, there were far more of them than there are top-level arguments because each proof about a given structure required a top-level argument and multiple property-verification arguments (for example, demonstrating that a structure is a ring requires a top-level argument that coordinates six property-verification arguments). Below is Dr. Tripp's written proof that the kernel of a ring homomorphism is an ideal:

K = {m| f(m) = f(0)}
We already know this is an ideal!
$f(0_R) = 0_S,\ 0_R \in K$
$\forall a,b \in K,\ f(a+b) = f(a) + f(b) = 0_S + 0_S = 0_S$
$a + b \in K$
$f(ab) = f(a)f(b) = 0_S f(b) = 0_S$
$f(ra) = f(r)f(a) = f(r)0_S = 0_S,\ ra,ar \in K$
[*Dr. Tripp erased the previous work and began writing on a clean board*]
$\forall a \in K,\ f(-a) = -f(a) = -0 = 0$
K is an ideal. [*Then underlined*]

There are five separate property-verification arguments represented in Dr. Tripp's proof: additive identity and non-empty (done together), additive closure, multiplicative closure, the ideal condition, and the existence of additive inverses. The additive identity proof consists of only two clauses: the data, "$f(0_R) = 0_S$;" an unwritten warrant, "We have proved that a homomorphism maps the identity to the identity;" an unwritten backing, "0_R is the additive identity element of *R*;" and a written conclusion, "$0_R \in K$," which corresponds to a stated but unwritten conclusion, "So, it's got something in it." This argument is diagramed in Figure 2.

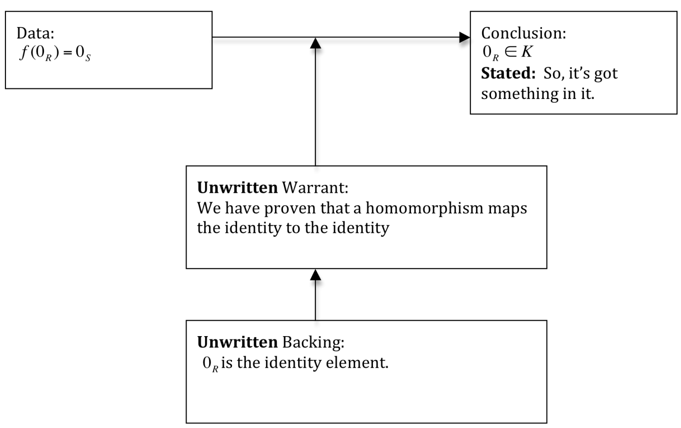


Figure 2. Dr. Tripp's non-empty argument.

Her subsequent arguments had a similar relationship between written and stated elements. In her proof that the set K is closed under addition, Dr. Tripp wrote:

$$\forall a,b \in K,\ f(a+b) = f(a) + f(b) = 0_S + 0_S = 0_S$$
$$a + b \in K$$

The statement $\forall a,b \in K,\ f(a+b) = 0_S$ is data, $f(a+b) = f(a) + f(b) = 0_S + 0_S = 0_S$ is a written warrant, and $a + b \in K$ is the written conclusion. Dr. Tripp did not write any backing for her warrant, but while writing out the proof, engaged in the following dialogue with a student:

Dr. Tripp: Ok, so, let's look at what *f* does to *a* + *b*. So, S, what can I say about *f*(*a* + *b*)?

> **S**: It equals *f*(*a*) plus *f*(*b*).
> **Dr. Tripp**: Is there anything I know about *f*(*a*) now?
> **S**: It equals zero.
> **Dr. Tripp**: Great, and *f*(*b*)? And what do I know about zero plus zero? That's zero. Great. So, $a + b$ meets the condition it needs to be in K. [pause] So, that's the property of f preserving addition that we just used, and that gives us that the kernel is closed under addition.

That is, Dr. Tripp stated, but did not write, two different backing statements for the written warrant. She first explained why it is permissible to rewrite $f(a+b)$ as $f(a)+f(b)$ and explained why $f(a)$ and $f(b)$ were both equal to the additive identity element in the ring S. We diagram her argument below.

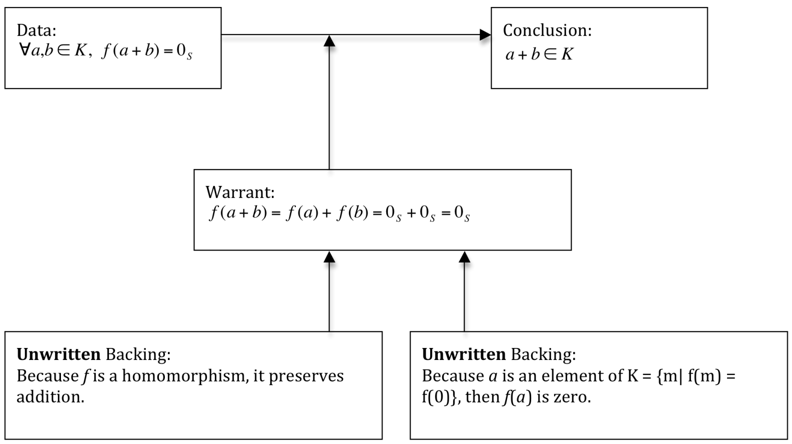


Figure 3. Dr. Tripp's closure argument

Dr. Tripp presented a written piece of data, conclusion and warrant. She stated, but did not write the backing that she provided. Analysis of the remainder of the property verification arguments that make up this proof show the same pattern; backing is stated, not written, while the rest of the argument is written.

To summarize Dr. Tripp's presentations in all of her proofs, data and conclusions are (almost) always written while the warrants were generally unwritten and included no qualifiers. However, the warrants were generally stated aloud during the classroom discussion, either by Dr. Tripp or by a student.

	Data	Warrant	Backing	Qualifier	Conclusion
Written	65	14	5	0	64

Table 1. Dr. Tripp's written argumentation

When we include her commentary in the argument analysis for the proofs presented above, we see the following:

	Data	Warrant	Backing	Qualifier	Conclusion
Written or stated	65	35	15	0	64

Table 2. Dr. Tripp's argumentation

That is, Dr. Tripp only infrequently wrote the warrant, and almost never wrote the backing, but modeled both fairly frequently in spoken form.

Given this pattern, we can expect to see similar trends in student proof-writing. We expect to see that students always or almost always wrote the data and conclusion, because Dr. Tripp always wrote those. But her mixed writing of warrants may not have provided her students a consistent model for their own work; thus, we can expect missed results in the written arguments given by the students. Although, we would expect the students to write warrants at a higher rate that Dr. Tripp did because of her frequency of stating them aloud. This is because Dr. Tripp shows a pattern of mixed written and spoken warrants when modeling proofs. Finally, we should expect that the students never, or almost never, write a backing or qualifier in their proofs, since Dr. Tripp never wrote either of these.

4.2 Analysis of student work

Taken together, the students made 33 different property arguments, which I analyzed using Toulmin's scheme. I will present the data on student arguments in three different sections, in the following order: additive property arguments, multiplicative property arguments, and the distributive property. I will show a representative sample of each and summarize the presence or absence of the different elements (data, warrant, backing, qualifiers, conclusion) in the students' written work. I will finish by summarizing the written elements of the students' arguments and comparing those with Dr. Tripp's modeled behavior.

4.2.1 The additive property arguments

Four of the students wrote proofs of the additive properties of closure, identity, inverses, and of the commutative and associate properties of all elements. When I analyzed these arguments using Toulmin's model, there were all essentially the same. The argument was given by the following relationship:

Data: Addition on the integers fulfills this property

Warrant: The operation $\oplus$ on this ring is defined as standard arithmetic on the integers.

Conclusion: R fulfills this property.

I provide an example of such an argument below to demonstrate that the ring has an additive identity element.

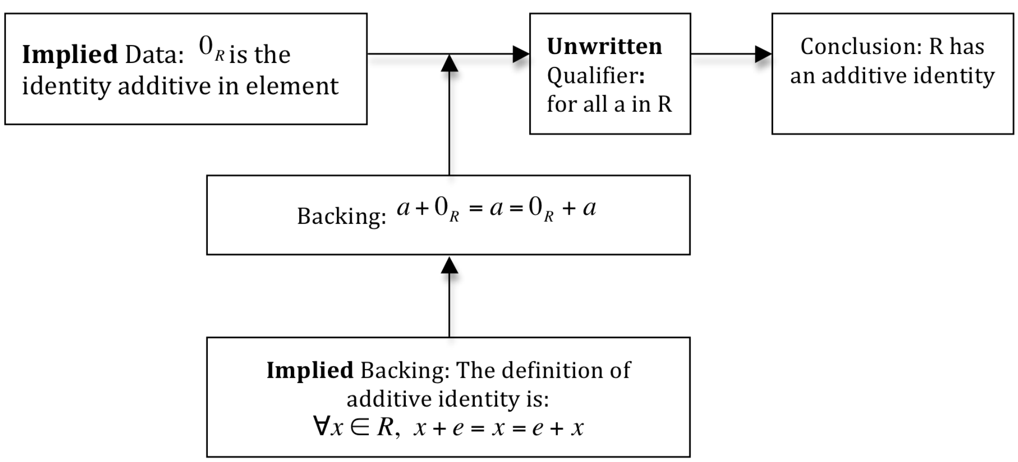


Figure 4. An example of a student argument about additive identity

This argument structure may be understood as characteristic of their work on addition properties. One student went so far as to summarize all the property arguments about addition at one time using the structure above. This student did not add any qualifiers and did not give any explicit backing. The table below summarizes the students' work on their proofs of additive properties:

	Data	Warrant	Backing	Qualifier	Conclusion
Stated	19	4	1		19
Unstated but implied	1	16			1

Table 3. Student's additive argumentation

In the proofs of these additive structures the students almost always stated the data and conclusion and only stated a warrant four out of sixteen times, a 25% rate. This low rate may be due to the fact that the students felt the additive properties of the integers were nearly self-evident by the middle of the semester (after 8 weeks spent studying rings) and thus, did not require a warrant or backing. The students typically wrote lengthier arguments in proofs of the multiplicative properties.

4.2.2 The multiplicative property arguments

The assessment required only that the students demonstrate that the defined structure was a ring, and, as a result, the students only attempted arguments that the defined multiplication was closed and associative. Because there were fewer properties to verify, there were fewer arguments to analyze. Moreover, possibly because the operation was less familiar, there was more variation in the structure of the arguments that the students wrote. For example, the students' arguments were more likely to state a warrant. One argument given was invalid. Multiple students included stated backing, and one student used a qualifier statement.

Let us begin by considering an argument for the closure of R under multiplication that includes a written warrant, unwritten backing, and a written qualifier. Jeff wrote:

(vi) If $a, b \in R$, then ab is either a or b, depending on which is larger. Either way, $ab \in R$ and so R is closed under multiplication.

A diagram of Jeff's argument using Toulmin's model:

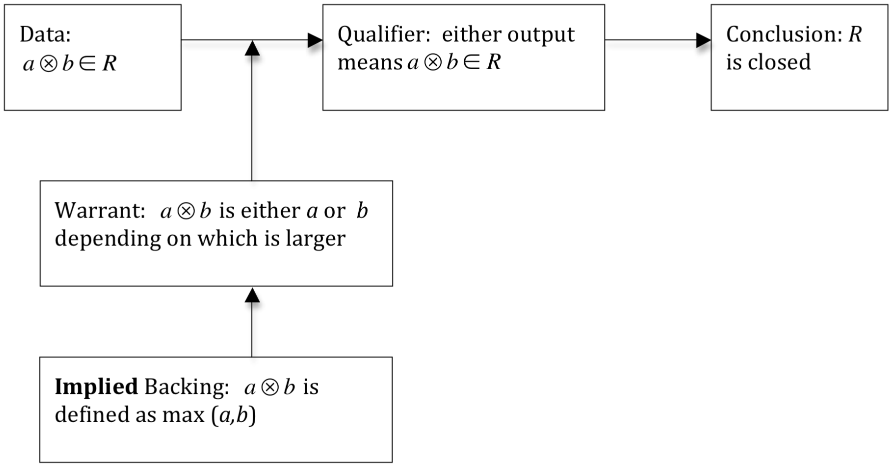


Figure 5. Jeff's argument for closure under multiplication

Aside from the written qualifier, this argument is illustrative of the closure arguments that the students wrote.

In all three of the arguments we have seen so far the student wrote at least one warrant. The table below summarizes the number of times that the students left unstated, but implied, different aspects of the argument about multiplicative properties.

	Data	Warrant	Backing	Qualifier	Conclusion
Stated	8	6		1	6
Needed but missing		2			

Table 4. Student's multiplicative property argumentation

This table suggests that Nathan, Jeff, and Aurora's work was reasonably representative of the class in terms of the elements of the argument that they wrote or left implicit (excepting Jeff's inclusion of a qualifier). In the examples above, the students wrote the data, the conclusion, and the warrants, but not the backing. This is reflected in the students' written work where we see that, generally speaking, they wrote the data, the warrant and the conclusion, but not the backing. This mirrored their work with the additive properties. The students wrote warrants at a higher rate than for the additive property arguments. This difference may result from the difference in the students' familiarity with the operations; the first operation of the proposed structure is very familiar to the students, standard integer addition. The second operation, the maximum operation, is less familiar to the students and therefore, they may have believed that this unfamiliarity meant that they should include more detail in their written proof.

4.3 Summary:

All of the student data on property-verification arguments is aggregated in the table below, without reference to the validity of the student's claim (one student made an invalid claim about distribution).

	Data	Warrant	Backing	Qualifier	Conclusion
Stated	32	14	2	1	30

Table 5. Student's argumentation

A few observations can be made. First, the students almost universally wrote both the data and conclusion of their arguments. They were much more varied when writing out (or implying) warrants in their work; of the 33 written arguments, only 14 included a written warrant. Second, the students almost never wrote out a backing or a qualifier, writing only two backings and one qualifier in 33 arguments. Yet, there was a difference in the rate at which the students wrote details between the purely additive arguments and those that involved the other operation the, maximum operation. The students wrote more details about the less familiar maximum operation in their proofs.

In sum, we see exactly what we expected from the students given Dr. Tripp's modeling of written proofs. She always wrote the data and conclusion, hardly even wrote a warrant (though she did say them along approximately half the time), and never wrote a backing or a qualifier. Clearly, in this case, the students performed the same way: they consistently wrote the data and conclusion, wrote a warrant in just under half of their arguments, and very rarely (6% of the time) wrote backing or a qualifier (3% of the time). In other words, the students seem to have appropriated Dr. Tripp's proof-writing behavior; they gave similar levels of argument detail in their written work.

5. Significance and directions for further research

This article makes three significant contributions. First, it shows one way to draw upon Toulmin's (1969) model of argumentation to analyze proof-presentation in undergraduate courses, including traditionally taught abstract algebra courses. This process required analysis of what the instructor wrote and of the classroom dialogue surrounding the proof-presentation. In both cases, Toulmin's model helped explain the relationship between the instructor's written proof and the classroom dialogue that she led. When Dr. Tripp's presentation is understood to model the type of mathematical behavior that she wants her students to demonstrate, we may conclude that she wants her students to always be able to articulate the data, warrant and conclusion of an argument. She also modeled use of backing and qualifier statements by stating them aloud, rather than writing them down. Dr. Tripp also wrote out warrants for her data slightly less than half of the time, but always spoke them. Thus, I demonstrate through my analysis that instructors may be inconsistent in the level of detail that they include in written arguments.

Second, I analyzed student work using Toulmin's (1969) model for argumentation. Taken as a whole, the students' collective proof-writing was fairly consistent. They almost always wrote the data and conclusion and avoided any written backing or qualifiers. Moreover, for the additive property verification tasks, the students essentially avoided writing out warrants, whereas on the multiplicative and distributive property verification tasks, they generally wrote the warrants.

Finally, I showed a link between Dr. Tripp's modeling of mathematical proof-writing and the students' demonstrated proof-writing. When I analyzed the students' work on property-verification arguments, I found that they wrote using a level of detail similar to that modeled by their instructor. The students were observed to write data, conclusions, warrants, backing and qualifiers at nearly the same rate that had been observed in Dr. Tripp's modeling of proof-writing. Furthermore, as noted above, Dr. Tripp was inconsistent in writing versus speaking aloud warrants, and we see that the students were similarly inconsistent, writing a warrant in slightly less than half, 42%, of all of their arguments. That is, the students almost perfectly demonstrated that they had appropriated Dr. Tripp's modeled proof-writing in terms of the level of detail that they included in their written work.

This research immediately suggests two future directions for research, both directed towards better understanding the development of students' mathematical proficiency. First, the use of Toulmin's framework to analyze teaching was helpful in making sense of some aspects of Dr. Tripp's writing and her classroom dialogue, but cannot explain all aspects of her modeling of proof-writing. We need significant research that studies the teaching of proof-writing (Harel & Sowder, 2007; Harel & Fuller, 2009), and, in particular, lecture-based teaching of proof-writing. Furthermore, we need new theoretical lenses to make sense of what lecture-based teachers are doing in classes that provide a means to explain student mathematical proficiency. In this vein, it is helpful to distinguish between introduction to proof courses and a proof-based content courses. A proof-based content course has the dual goals of improving the students' ability to write and critique proofs as well as increase their proficiency with the particular mathematical content of the course. In contrast, an introduction to proof course has one goal: to teach students the logic and mechanics of writing proofs. Thus, theoretical models focused solely on the teaching of proof-writing or proof-comprehension will be insufficient if they cannot help researchers account for instructors' attempts to teach content, as well.

Thinking of Dr. Tripp as modeling behavior that she wanted students to adopt was helpful in understanding her classroom behavior. Thus the idea that teachers consciously model appropriate mathematical behavior for their students is worth pursuing as a means of making sense of lecture-based undergraduate courses. For example, we should explore how instructors model other fundamental mathematical skills for their students, such as definitions and examples, organization and connection of knowledge, and abstraction and generalization from examples, to name but a few. It is important to note that this list is suggested by the habits that mathematicians believe are important in the practice of mathematics and may not completely capture the list of behaviors that instructors actually model for their students. It almost certainly does not explain what is presented as the *work* of doing mathematics—the research methods of mathematicians, their habits and practices. Significantly more research is needed to explain traditional lecture-based instruction of proof-based content courses in general, and abstract algebra courses in particular.

Lastly, given a goal of improving student learning, this line of analysis offers one possible avenue for mathematics educators to work with mathematicians to make minor modifications of their teaching that may lead to increased student success. In this study I reported seemingly inconsistent behavior on the part of the instructor, which was linked to inconsistent behavior on the part of the students. We might suggest that classroom professors make explicit statements, such as, "Listen to the types of questions I ask you and ask myself. These are the kinds of things *you* should be asking while you write proofs." Or, "To decide when you should write down the answers to these questions you should …" This type of meta-dialogue could lead the students to a better understanding of when warrants and backing must be explicitly stated in a proof, as well as decrease the writing of invalid or incorrect proofs. Given that Dr. Tripp's class, and by extension, many other lecture-based classes, already include significant conversation and dialogue, this addition to the classroom discussion seems a relatively small change that may have a more significant effect on students' learning than a more difficult switch to an inquiry-oriented pedagogy.

References

Alcocok, L. (2009). Mathematicians' perspectives on the teaching and learning of proof. *CBMS Issues in Mathematics Education, 16*, 73-100.

Douglas, R., (Ed.), (1986). *The Importance of Calculus in Core Mathematics Toward a Lean and Lively Calculus*. Washington, DC: Mathematical Association of America.

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics, 40*(1), 85-109.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics, 27*(3), 267-305.

Fukawa-Connelly, T. (2010, February). *Modeling Mathematical Behaviors; Making Sense of Traditional Teachers of Advanced Mathematics Courses Pedagogical Moves*. Raleigh, NC: Conference on Research in Undergraduate Mathematics Education.

Harel, G. & Fuller, E. (2009). Contributions toward perspectives on learning and teaching proof. In D. Stylianou, M. Blanton, and E. Knuth (Eds.), *Teaching and learning proof across the grades; A K-16 perspective.* New York, NY: Routledge.

Harel, G. & Sowder, L. (2007). Towards a comprehensive perspective on proof. In F.

Lester (Ed.), *Second handbook of research on mathematical teaching and learning*. Washington, DC: NCTM.

Hungerford, T. W. (1997). *Abstract algebra-An introduction* (2nd ed.). Florence, KY: Brooks/Cole.

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modeling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66*, 3–21.

Knipping, C. (2008). A method for revealing structures of argumentation in classroom proving processes. *ZDM, 40* (3), 427-441.

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb, & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Krummheuer, G., (2007). Argumentation and participation in the primary mathematics classroom: Two episodes and related theoretical abductions. *Journal of Mathematical Behavior 26*, 60–82.

Leron, U., Hazzan, O., & Zazkis, R. (1995). Learning group isomorphism: A crossroads of many concepts. *Educational studies in mathematics, 29*(2), 153-174.

Mathematical Sciences Education Board (MSEB). (1991). *Moving beyond myths: Revitalizing undergraduate mathematics*. Washington DC: National Academies Press.

McClain, K., & Cobb, P. (2001). The development of sociomathematical norms in one first-grade classroom. *Journal for Research in Mathematics Education. 32*(3), 236-266.

National Science Foundation. (1992). *America's academic future: A report of the Presidential Young Investigator Colloquium on U.S. engineering, mathematics, and science education for the year 1010 and beyond*. Washington DC: Author.

Pemberton, J., Devaney, R., Hilborn, R., Moses, Y., Neuhauser, C., Taylor, T., Brady T., & Williams, R., (2004). *Undergraduate education in the mathematical and physical sciences: Interim report of the Joint Subcommittee of the NSF Education and Human Resources and Mathematical and Physical Sciences Advisory Committees*. Washington DC: National Science Foundation. Retrieved from http://www.nsf.gov/attachments/102806/public/JSACInterimReportEHR-MPSSpring2004.pdf.

Rasmussen, C. & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics into instruction. *Journal for Research in Mathematics Education*, *37*, 388-420.

Richarson, V. & Placier, P., (2001). Teacher Change. In Richardson, V. (Ed), *Handbook of Research on Teaching*, (pp. 905-947). Washington DC: American Educational Research Association.

Speer, N. M. & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. *Journal for Research in Mathematics Education 40*(5), 530-562.

Speer, N., Smith, J., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, *29*, 99–114.

Sowder, J. (2007) The mathematical education and development of teachers. In Lester, F. (Ed.), *Second Handbook of research on mathematics teaching and learning* (pp. 157-224.) . Reston, VA: NCTM.

Steen, L. A. (Ed.), (1987). *Calculus for a new century.* Washington, DC: The Mathematical Association of America.

Toulmin, S. (1969). *The uses of arguments.* Cambridge: Cambridge University Press.

Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction, 24*(1), 1-71.

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

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior, 23*, 115-133.

Weber, K., Maher, C., Powell, A., & Lee, H. (2008). Learning opportunities from group discussions: Warrants become the objects of debate. *Educational Studies in Mathematics, 68*, 247–261

Wu, H. (1999). The joy of lecturing—With a critique of the romantic tradition of education writing. In S. G. Krantz (Ed.), *How to teach mathematics* (pp. 261-271). Providence, RI: American Mathematical Society.

Yackel, E. (2001). Explanation, justification, and argumentation in mathematics classrooms. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference for the International Group of Psychology of Mathematics Education, 1*, 9–24, Utrecht, the Netherlands.

COMPREHENDING LERON'S STRUCTURED PROOFS

Evan Fuller
Montclair State University
fullere@mail.montclair.edu

Juan Pablo Mejia Ramos, Keith Weber, Aron Samkoff, Kathryn Rhoads,
Dhun Doongaji, Kristen Lew
Rutgers University

In a widely cited paper, Leron (1983) proposed presenting proofs in a novel format that he called "structured proofs" and suggested that this format improves comprehension. In a qualitative study, we found that participants had difficulty with this format for several reasons, including a lack of familiarity with the format and the requirement on the reader to jump around between different sections of the proof. In a larger quantitative study, we found that students reading a structured proof were better than students reading a linear proof at identifying a good summary of the proof, but performed slightly (but not statistically reliably) worse on questions concerning justifications within the proof, transferring the ideas from the proof to another setting, and illustrating the ideas of the proof using examples.

Key words: Proof; Proof comprehension; Structured Proofs

1. Introduction

1.1. Proof presentation and comprehension in advanced mathematics courses

Advanced mathematics courses—that is, tertiary proof-oriented mathematics courses for mathematics majors—are typically taught in a lecture format, a significant portion of which consists of the professor presenting proofs of theorems to his or her students. Based on her observations of the lectures of three mathematics professors, Mills (2011) reported that roughly half the lecture time was spent on proof presentation. Numerous mathematicians and mathematics educators have commented that mathematics is typically presented to students in a definition-theorem-proof format (e.g., Davis & Hersh, 1981; Dreyfus, 1991; Thurston, 1994; Weber, 2004).

Presumably an assumption behind this pedagogical practice is that students can learn mathematics by reading and studying the proofs that their professors present or that appear in textbooks and notes. Although there have been few systematic studies on undergraduates' comprehension of proofs (as noted by Mejia-Ramos & Inglis, 2009), there are many who question whether this assumption is accurate. Both mathematicians and mathematics educators have remarked that students are generally confused by the content of a formal proof (e.g., Alcock, 2010; Davis & Hersh, 1981; Hersh, 1993; Leron & Dubinsky, 1995; Rowland, 2001; Thurston, 1994), with students reporting that the proofs they see are pedantic or pointless (e.g., Porteous, 1986). Further, empirical studies demonstrate that undergraduates have difficulty distinguishing between valid proofs and invalid arguments (Selden & Selden, 2003; Weber, 2010). There are, of course, many potential reasons that undergraduates gain little from the proofs that they read. However many researchers suggest that some student difficulties with proof are due to factors inherent in the way formal proofs are written.

Griffiths (2000) defines a proof as "a formal and logical line of reasoning that begins with a set of axioms and moves through logical steps to a conclusion" (p. 3). These types of proofs usually proceed linearly, incorporate logical syntax, and make little use of informal representations of the relevant mathematical concepts, such as diagrams and examples. Mathematicians and mathematics educators argue that this type of presentation might hinder proof comprehension for several reasons. The linear nature of proof presentation can prevent students from seeing the structure of the proof or the overarching method being applied in the proof, making the ideas of the proof appear to be unmotivated and mysterious (Davis & Hersh, 1981; Leron, 1983). The use of logical syntax and domain-specific jargon can be intimidating to students and mathematicians alike (Davis & Hersh, 1981; Hersh, 1993; Thurston, 1994). Finally, the formal presentation of proofs masks the intuitive mathematical ideas and models that were used to produce the proof (Dreyfus, 1991; Thurston, 1994).

To address this situation, several researchers have suggested alternative methods of proof presentation. These include the use of generic proofs (Rowland, 2001), informal visual arguments in lieu of formal proofs (Hersh, 1993), e-proofs (Alcock, 2009), and structured proofs (Leron, 1983). Although each of these approaches is interesting and potentially valuable, there is to date little to no empirical evidence that any of these approaches improve proof comprehension. Indeed, Roy, Alcock, and Inglis (2010) found that students who studied a proof using Alcock's e-proofs (Alcock, 2009) performed significantly worse on a comprehension post-test than students who observed the same proof in a lecture. We are not aware of any other studies that examine the effects of these novel proof presentations on students' proof comprehension.

1. 2. Leron's structured proofs

In this study, we examine the ways in which Leron's (1983) structured proofs can improve students' comprehension of a proof. Leron proposed a novel way to present proofs in terms of levels. The highest level (Level 1) provides a summary of the main ideas of the proof without providing detail on how these main ideas will be carried out. The next level (Level 2) provides a summary of how each of the main ideas will be implemented. Successively lower levels fill in the details of the implementation of higher levels of the proof. An additional feature in some structured proofs is an "elevator" between levels that provides a rationale for why the proof is proceeding the way that it is. Leron illustrated the nature of structured proofs by comparing a linear and a structured proof of the following claim: "There are infinitely many triadic primes". We used these proofs in our studies (nearly verbatim from Leron's article) and they are presented in the Appendix of this paper. The reader can note that in this proof, Level 1 lays out a summary of the three main aims of the proof, while the "elevator" provides the motivation for introducing and defining the variable M in a particular way.

There are several theoretical benefits for using structured proofs in lieu of traditional linear proofs. The format provides the reader with a summary of the proof and enables the reader to grasp the main ideas of the proof without getting lost in its logical details. However, no information is lost as this format still enables the reader to study or verify these logical details if he or she desires to do so. Finally, the higher-level structure and the "elevator" commentary make clear the reasoning behind some of the choices made in the proof that might otherwise seem arbitrary.

Numerous mathematics educators approvingly cite Leron's structured proofs as a potential way to improve students' proof comprehension (e.g., Alibert & Thomas, 1991; CadwalladerOlsker, 2011; Hanna, 1990; Hersh, 1993; Selden & Selden, 2003, 2008). In a book chapter providing research-based pedagogical suggestions to mathematicians, Selden and Selden (2008) recommended Leron's structured proofs as a means of helping students understand the proofs that they read and learn about the process of writing proofs. Mamona-Downs and Downs (2002) suggested that Leron's structured proofs may be influencing the ways in which proofs are currently written in advanced mathematics textbooks. Melis (1994) wrote that "Uri Leron shows how proofs are better comprehensible by structuring them into different levels" (p. 2), implying that the claim that structured proofs improve comprehension is not a theoretical hypothesis but rather an established fact.

Despite the theoretical promise of Leron's structured proofs and the esteem in which Leron's suggestion is held by some in the mathematical community, we are not aware of any empirical studies that suggest that structured proofs actually improve comprehension. Based on their observations of three students reading a structured proof using hyperlinks in a computer environment, Cairns and Gow (2003) noticed several difficulties that students had with this format. The novelty of the format did not map easily to the way that students were used to reading proofs (in a linear fashion), the students had difficulty evaluating their progress while reading the proof, and relevant information within the proof was often hidden or spaced far apart. However, given that Cairns and Gow investigated students' reading of proofs using hyperlinks (rather than as traditional text), and that the proofs were presented using an approach that was similar but not identical to Leron's (specifically, the approach of Lamport, 1995), the relevance of these findings to Leron's structured proofs should be viewed as suggestive at best.

1. 3. Research questions

The general aim of this paper is to examine the efficacy of Leron's structured proofs—to what extent do structured proofs improve comprehension, and what difficulties do students experience when reading them? However, we first emphasize the limitations and qualifications inherent in this type of research. Schoenfeld (2000) warns that questions such as "does group work improve mathematical learning?" are not well-formed. Before attempting to address this question, one would need to specify what type of group work was being implemented, what context was being studied, and what constituted learning. Even here, a single study or series of studies would not produce conclusive results as one would need to consider the quality of the implementation of the instruction, among other factors.

The pedagogical suggestion of structuring proofs can be implemented in a variety of ways. Students could be asked to read structured proofs as text in a short period of time, a student could be asked to study a structured proof overnight, or a professor could present a structured proof in lecture for his or her students. Structured proofs can even serve as an overarching paradigm for instruction, where many proof presentations involve structured proofs and students are involved in the structuring of some proofs themselves (indeed, a former student of Leron informed us that this was what he did in his own classroom). In this paper, we compare students' comprehension of a proof shortly after reading a linear or structured version of the same proof. However, any learning benefits that we did not observe in this study might occur if the structured proofs were used in a different way (and vice versa).

The two specific research questions we investigate are:

- What specific features of a structured proof help or hinder proof comprehension?
- To what extent do students who read a structured proof display a greater ability to (1) summarize the proof, (2) transfer the ideas of the proof to prove a different theorem, (3) illustrate the ideas of the proof with a specific example or diagram, or (4) see how particular assertions within the proof were justified, than students who read the same proof as a traditional linear proof?

Even here, as we only tested these ideas with two proofs, we do not view our answers as conclusive, but rather as the start of a conversation on the benefits and limitations of structured proofs.

2. Theoretical perspective

The assessment items that we use in this paper are based on our assessment model for proof comprehension (Mejia-Ramos et al, 2010). Based on interviews with mathematicians and our survey of the mathematics education literature, we developed seven different ways to assess students understanding of a proof. However, only four were utilized in the quantitative study that we focus on here:

- *Justification of claims*. Questions of this type assess whether students can justify how a new statement in a proof follows from previous assertions as well as how a particular statement is used to justify subsequent statements within the proof.
- *Summary*. Questions of this type assess whether students can recognize the overarching ideas used in the proof.
- *Transfer*. Questions of this type assess whether students can apply the ideas used in the proof in a new setting to prove a new theorem.
- *Illustration with examples*. Questions of this type assess examine whether students can relate the chain of argumentation in a proof to a particular example (i.e., transform the formal proof into a generic proof) or a relevant diagram.

The rationale for developing all seven categories, as well as a scheme for generating questions of each type, are described in detail by Mejia-Ramos et al (2010, submitted). Leron's structured proofs seem theoretically designed to help with the summary portion of this model (since Level 1 is essentially a summary of the proof) and transfer (since Level 1 and the between-levels elevator provide insight into the method being applied).

We note that our assessment model for proof comprehension is not designed to be hierarchical, but rather is a means to assess separate dimensions of proof comprehension. We can imagine one type of proof presentation helping students with some of the dimensions of this model while giving no benefit or hindering students with others.

3. Study 1: Qualitative interview study

3. 1. Rationale

We first conducted a set of interview studies in which students were videotaped while reading linear or structured proofs, answering questions about those proofs, and responding to open-ended questions about how they felt about the structured proof presentation. The purpose of these interviews was twofold: First, we wanted to examine how students read and reacted to structured proofs to gain insight into how this type of proof presentation might help or hinder students' comprehension of a proof. Second, we wanted to use students' behavior during our study to improve the quality of the proofs and assessment questions that we used during the quantitative study that we report in the next section.

3. 2. Methods

Participants. We recruited three groups of six participants to take part in this study. Participants were paid a fee for their participation. We recruited Groups A and B from math major courses at a large state university in the northeast United States in the beginning of the Fall 2010 semester.

After interviewing the students in Groups A and B, it became clear that some of them were deeply confused about the nature of structured proofs. We believed one cause of their difficulty was that they were not given a description about the nature and purposes of these structured proofs. At the end of the Fall 2010 semester, we recruited six more students and asked them to read two structured proofs, but only after giving them instructions about the nature of a structured proof (these instructions appear in the Appendix).

Materials. Two proofs were used for this study. We called the first proof the "Only Zero" proof, since it used a calculus-based argument to prove that 0 was the only solution to a given equation. The rationale for using the Only Zero proof is that we believed the content of the proof would be accessible to second- and third-year mathematics undergraduates and the proof could be effectively structured. The second proof, which we call the "Triadic Primes" proof, establishes that there are infinitely many triadic primes (i.e. primes congruent to 3 modulo 4). The linear and structured proofs we presented to participants in this study were taken nearly verbatim from Leron's (1983) original article on structured proofs. These proofs are both included in the Appendix of this paper.

Procedure. During each semi-structured interview, participants were given a version of the Only Zero proof and asked to read this proof until they felt that they understood it. The participants were informed that after they had studied the proof, the proof would be taken away from them and they would have to answer questions about what they had just read.

When participants finished reading the proof, they were asked to say, on a scale of one through five, how well they felt they understood the proof. They were then asked the questions about the proof, one at a time, and were given a sheet of paper with the question written on it. In interviews with Groups A and B, participants were given a set of multiple choice questions following the open-ended questions, most of which were similar to the questions they had just answered. As we explain shortly, one goal of having students answer these multiple choice questions was so we could refine the questions for our quantitative study. If participants read a structured version of the proof, they were asked how they felt about this new format of the proof, whether they liked this new format, and what they thought the strengths and weaknesses of this format were. This entire process was then repeated with the Triadic Primes proof. Since we found that we had sufficient data from Groups A and B in order to refine the multiple choice questions, Group C was given only open-ended questions in the interviews.

Group A received a linear version of the Only Zero proof and a structured version of the Triadic Primes proof. Group B received a structured version of the Triadic Primes proof and a linear version of the Only Zero proof. Group C first read the instructions describing the nature of structured proofs that are presented in the Appendix and then read structured versions of both the Only Zero and Triadic Primes proofs. Interview lengths ranged from 60 to 150 minutes.

Analysis. Our analysis consisted of two phases. First we sought to identify attributes of structured proofs that may have aided or hindered the participants' comprehension of the proof. In an initial pass through the data, we noted three sources of difficulty that participants had with structured proofs: Participants expressed displeasure that the proof seemed to "jump around", they had misconceptions about the nature or purpose of structured proofs, and they failed to understand the high-level summary expressed in the top level of the structured proof.

In the second phase of analysis, we aimed to improve the readability of the structured proofs and the quality of the questions for the large-scale quantitative study that we conducted subsequently. To this end, we noted any common confusions that participants had about the proof that was not related to the format of the proof and sought to revise the proofs to avoid these errors. For instance, many participants had trouble inferring what "triadic prime" meant from the definition of triadic number. This was explicitly explained in the proof for the follow-up study. We also examined any instances in which participants answered the open-ended questions incorrectly but answered the multiple choice questions correctly, and vice versa, and formulated new multiple choice questions that more accurately measured participants' understandings of the proofs that they read.

3. 3. Results

We noticed three common difficulties that participants had when reading structured proofs that may have inhibited their comprehension of the proof.

Jumping around: Fourteen of the eighteen participants had difficulty reading the structured proofs because they felt the proof jumped around between the main ideas of the proof, including five of the six participants who received instructions on structured proofs. We provide three representative comments below, including two from students in group C (C1 and C3) who received instructions:

> B3: I'm not really sure the reason why it couldn't be just one flowing proof. It kind of bounced from parts, like from 2b, it would go to 3b, or whatever it is.
>
> I: Do you think this type of format increased your understanding of the proof, decreased your understanding, or neither?
>
> B3: Probably decreased it slightly, just because it was moving around with the statements a lot, jumping from the middle of the page to the bottom of the page. Things weren't flowing one after another, which made it a little more confusing[…] it just didn't allow for one thing to flow to the next in a way that was easy to read.
>
> I: In what ways if any did this format hinder your understanding of the proof?
>
> B3: The same way. It just bounced around too much.

> I: What did you think of the format?
>
> C1: I kind of like it because in a sense it's kind of how I would break down proofs sometimes. Because when I do a proof, if I think of something, I'll do a little scratch work over here and then more scratch work over here, and I had a sense that it was like

that. But since there was just so much stuff and you had to keep going back and drawing arrows, that was the only confusing part […] Just because someone says something is true doesn't necessarily mean it's true, so I kind of want to show for myself that it's true, so that's what it did. But just because the stuff was all over the place and there was like in three places like we support the claim in 2c, but by the time I got down here I forgot what 2c was, so I had to keep going back and drawing. [This participant drew arrows between the assertions in Level 1 and the sub-proofs in Level 2].

C3: But if you were to read it like that like I did the first time, you're jumping around everywhere. Because you're saying level two, but I don't really read level two because I see more text down here. So I figure let me read this first. [the rest of level 1]. But then it gets very confusing. Because for example I read 2c, and I go back to 3a but then they're talking about the claim in 2b. But I don't remember the claim in 2b, so I have to go back and read 2b properly. But then I have to remember how does this relate to what the goals were? And go back here, and I'm going all over the place.

The "jumping around" is, of course, a crucial feature of structured proofs, but the students comments reveal a limitation of this feature—specifically, participants have to keep several different ideas in their heads as they read these proofs, which likely puts a strain on their working memories. This difficulty that the participants experienced is related to Cairns and Gow's (2003) that this genre of proofs does not keep relevant information near each other.

Failure to understand the high-level summary. For the structured Triadic Primes proof, many participants could not understand the main ideas expressed in Level 1 of the proof. When asked to provide a summary of the structured Triadic Prime proof, six of the twelve participants who read this proof either provided no response or a response that was fundamentally flawed. For instance, one participant, C1, began her summary by defining $3p_2p_3\ldots p_n + 1$ as a triadic prime. Further, participants who did provide sensible summaries claimed that this was not due to Level 1, which they found confusing, but rather from what they gained from studying the proof itself. Consider A1's comments below:

A1: Reading the proof for the first time I got a little confused, because we're talking about M as if it had already been constructed but then I was starting to think 'Oh well what's M?', and I realized that it hadn't actually been constructed yet. You had to get down here (pointing to a lower level) and I was trying to think about it and I was like, 'Wait! What's M? Am I missing something?' And then when I got further down I understood.

From our perspective, a substantial benefit of including Level 1 in a structured proof is to provide a framework or schema to integrate the chain of deductions that follow by providing them with purpose. However, if participants fail to adequately understand Level 1, they will either be trying to integrate the ideas from Level 2 into an inappropriate framework, which could lead to confusion, or they might forget the ideas contained in Level 1, leading to a strain on their working memory and a dissatisfaction that the proof was jumping around. This concern is conveyed nicely in the comments from C2.

C2: [Structured proofs] were a disaster, to be blunt. That proof using Rolle's theorem and using derivatives and stuff. I took analysis so we did proofs like that for an entire semester. So when I was faced with that proof, I already knew where it was going, I had seen things like it before, it was very familiar to me. So even though the style of the proof

> wasn't exactly my cup of tea I was still able to absorb a lot of the information because I was familiar with saying ok you know we're using a derivative argument, we're using a Rolle's theorem argument and what not. But for this proof [the Triadic Primes Proof], like I said the only proof I've ever seen in that form was the proof that there were infinitely many primes and if you asked me to write that down right now, I probably wouldn't be able to, to be honest. Although I think it's very elegant, it's not a proof skeleton or a proof form that I use very often, so although I knew what the proof was getting at, I wasn't able to absorb it as well simply because it's not something I see very often.

C2 argues that he was able to understand the structured proof of the Only Zero theorem in spite of the structured proof format because he had familiarity with the type of argumentation used in the proof. He lacked similar experience in number theory and consequently could not use the proof skeleton in Level 1 to "absorb" the details of the proof. Cairns and Gow (2003) note that structured proofs might be ineffective for students who do not possess a strong enough understanding of the material to see why the proof is structured the way that it is.

Failure to understand the nature of structured proofs. Participants in our study often held significant misconceptions about what a structured proof was and how a structured proof works. Three participants believed the structured proofs that they read were proofs by cases and one participant thought the structured proof was three different independent proofs of the same claim. Five participants expressed confusion regarding how the concluding sentence for why the theorem was proved occurred in Level 1, whereas the final sentence of the proof that appeared in Level 3 did not establish the theorem. For instance, when reading the structured triadic primes proof, one participant, C4, reached the end of the proof and said, "Is that it?", and turned the paper that the proof was written on over to look for more text. When told that was the proof in its entirety, C4 remarked, "I feel like we didn't prove anything." Later, when asked why he didn't understand the proof, C4 remarked:

> C4: I feel like we didn't prove anything […] I just didn't understand the ending at all […] At the end, showing that when you multiply each together you're going to get that kind of number [the end of the proof was a proof of the assertion that the product of monadic numbers was monadic], I'm going to say I'm really confused.

A main motivation behind structured proofs is that the proofs of the lemmas used in the proofs, which are less significant to the main line of argumentation, appear near the end of the proof to reflect their relative lack of importance. However, the participants who read these proofs did not appreciate this and expected the last line of the proof to conclude with the theorem being proven.

Most likely one contributing cause to this difficulty was the participants' lack of familiarity with this way of writing proofs and this would likely be mitigated if participants were given more experience with structured proofs. Indeed, eight participants remarked on the novelty of this format with some suggesting that their lack of familiarity with structured proofs may have limited comprehension. For example, in describing the structured proofs, B5 said, "It was kind of all over the place. I've never seen anything like it. [laughing] It was just bizarre." Cairns and Gow (2003) also noted that students' lack of familiarity with these types of proofs might pose a barrier to comprehension.

4. Study 2: Quantitative internet study

4. 1. Rationale for this study

In the previous section, we highlighted several potential attributes of structured proofs that might inhibit students' comprehension of a structured proof. However these data are limited for two reasons. First, the fact that students found some aspects of structured proofs to be problematic does not mean that structured proofs might not be beneficial for comprehension. It is possible that the other theoretical benefits of structured proofs more than compensate for the reported drawbacks. It is also plausible that the difficulties that students experience are not drawbacks at all; for instance, even though "jumping" between different sections of the proof was unpopular with our participants because it made the process of reading the proof longer and more difficult, perhaps this activity helped students make connections between ideas in the proof that they would have missed if they had read the proof in a linear format. Second, the sample of eighteen students reading structured proofs was relatively small, and only six students read structured proofs after being given instruction about the nature of structured proofs. Hence the findings from the qualitative interview study might not be generalizable.

The goal of the larger quantitative internet study was to search for systematic evidence with a larger number of students that structured proofs might help students summarize proofs, transfer the ideas of the proof to other settings, understand how the proof relates to a specific example or diagram, and see how particular statements within the proof were justified.

4. 2. Research methods

To maximize our sample size, we conducted the quantitative study recruiting mathematics majors using the internet. The validity and reliability of this type of studies have been extensively discussed in the research methods literature (e.g. Gosling et al., 2004; Reips, 2000). To deal with the common threats of validity, we employed the methodology of Inglis and Mejia-Ramos (2009) to discard participants who showed evidence from participating in the experiment twice (e.g., multiple submissions from the same IP address) or who had revisited earlier pages in the study.

Participants. We recruited 300 mathematics undergraduate students from 50 top-ranked mathematics departments in US universities. These students participated without payment and were recruited via an email from their departmental secretary. The email explained the purpose of the experiment and asked third and fourth year mathematics major/minor students to visit the experimental website if they wished to participate.

Materials. All materials used for this study are presented in the Appendix. For the internet study we used the two proofs employed in the qualitative interview study (i.e. the Only Zero proof and the Triadic Primes proof) with minor modifications. As mentioned earlier, we made these minor modifications in order to avoid confusion unrelated to the format of the proof. For instance, for the "triadic primes" proof, in addition to defining *triadic (monadic) numbers* right before the theorem and its proof, we added a statement clarifying the meaning of *triadic (monadic) primes*.

For each proof, we designed a proof comprehension test consisting of four (for the Triadic Primes proof) or five (for the Only Zero proof) items. As mentioned earlier, these assessment items were selected and improved from the questions used in the interview study.

Procedure. Depending upon the school they were at, participants were invited to read the Only Zero proof or the Triadic Primes proof. Once participants had loaded the experimental website, they were randomly assigned into one of three conditions: LIN where participants were presented with the linear version of the proof; S(no dir) where participants were presented with the structured version of the proof but with no directions; and S(w dir) where participants were presented with the structured proof after reading the instructions describing the nature and purpose of structured proofs (in the Appendix).

After reading the proof, participants were asked, "How well do you feel you understand this proof?" A five-point Likert scale was used to record each participant's reported level of understanding. Next, participants took the corresponding comprehension test. Each of the questions appeared in a new screen and the order in which they were displayed was randomized for each participant. Upon completion of the test, participants in a structured condition were asked to use a three-point Likert scale to report the extent to which they liked the format in which the proof had been presented.

4. 3. Results

We found no statistically significant differences between the performance on the items between the Structured-No Directions and Structured-Directions condition and, surprisingly, the Structured-No Directions did better than the Structured-Directions conditions on most items. Consequently, for the purposes of comparing the efficacy of structured proofs, we collapsed the Structured-No Directions and Structured-Directions groups into a single Structured group in what we report here. Participants' performance in the comprehension tests is presented in Table 1.

Table 1: Performance in linear versus structured conditions.

Only Zero proof

Condition	Summary Diagram	Transfer	Justification	Justification Question 1	Application to Question 2
Lin (N=67)	63%	61%	85%	85%	69%
Struct (N=135)	76%	53%	88%	80%	78%

Triadic Primes proof

Condition	Summary	Transfer		Application to Justification	Examples
Lin (N=33)	67%	45%	55%	91%	
Struct (N=65)	74%	38%	40%	63%	

Summary. The Structured Proof group did better on the summary questions than the Linear Proof group. For the Only Zero proof, the Structured Proof group answered the summary question correctly 76% of the time as compared to 63% for the Linear Group, a statistically reliable difference (c(1, 200)=4.095, p<0.05). For the Triadic Primes proof, the Structured Proof also did better on the summary question, although this difference was not statistically reliable (c(1, 96)<1). In structured proofs, Level 1 essentially provides a summary of the proof and, while some students may find it difficult to understand (see section 3), others may grasp it. The data demonstrates that some of the participants who read a structured proof were better able to retain this information and recognize a high-quality summary.

Transfer. We did not observe evidence that the Structured Proof group outperformed the Linear group on the transfer questions. On the contrary, although inconclusive, the evidence suggests a trend in the opposite direction. For both the Only Zero and Triadic Primes proofs, the Linear Group had a higher percentage of students who answered the transfer question correct, although neither of these results were statistically reliable (c(1, 200)=4.095, p=0.24, c(1, 96)<1, p=0.43). Consequently, the data provide no evidence that structured proofs improve students' ability to answer these types of transfer questions and suggestive evidence that they hinder students' abilities in this regard.

Justification. We did not observe evidence that the Structured group outperformed the Linear group on the justification questions. We found essentially no difference between participants' performance on the Justification questions for the Only Zero proof. For the Triadic Primes proof, the Linear group performed somewhat better than the Structured group (55% vs. 40%), but this difference was not statistically reliable (c(1, 96)=1.872, p=0.17). Hence, like the transfer questions, our results are inconclusive, but to the extent that they suggest anything, they favor the Linear group.

Illustration with examples. This was the only category in which we observed different findings on the Only Zero and Triadic Primes proofs. On the Triadic Primes proof, the Structured Group performed better than the Linear Group in recognizing that the ideas in the argument would not be applicable to a graphed function (78% vs. 69%), although this difference was not statistically reliable ((c(1, 200)=1.974, p=0.16). However, for the Triadic Primes proof, the Linear Group did substantially better on the question illustrating the ideas of the proof with a specific example (91% vs. 63%; c(1, 96)=9.252, p<0.01). In hindsight, it appears that evaluating whether the ideas of a proof *cannot* be applied to a specific diagram and actually applying the ideas of the proof to a specific example may be two very different skills.

In analyzing the conflicting results for illustration with examples, we suggest two factors. First, once M is chosen in the Triadic Primes proof, the linear version of the proof essentially offers a step-by-step recipe for reaching a contradiction with this M. For the structured proof, this information is dispersed throughout the proof and more emphasis is given to overarching methods and the motivation for choosing M. Reading a linear proof may help students with these types of questions because the step-by-step procedure is made explicit. Second, our post-hoc analysis reveals that when Leron (1983) originally posed the structured proof for the Triadic Primes Theorem, he did not just change the presentation of the linear proof, but he also changed the *content* of the linear proof as well. The linear version of the Triadic Primes proof reaches a contradiction by showing that M must be monadic and triadic while the structured version of the proof arrives at a contradiction by showing that M must have a triadic prime factor not included in the finite list of triadic primes. As a result, the assessment question that we asked the participants may have unintentionally favored the linear group, leading to the observed differences. Hence, despite the statistically reliable effect, we also view this result as inconclusive.

Participants' subjective evaluation of structured proofs. Detailed results on this topic will be presented in a forthcoming paper. For the purposes of brevity, we provide a summary here. There was remarkably little difference between the participants' condition and how well they understood the proof that they read. However, we found that participants in the structured proof with directions were about 50% more likely to view the structured proof format positively and 50% less likely to view the proof negatively. Hence, although providing directions did not improve students' professed understanding of a structured proof or their performance on the comprehension of the test, it did lead them to view structured proofs in a more positive light.

Finally we note that the correlations between participants' reported understandings of the proof and their cumulative comprehensive scores were negative, both for the Only Zero proof (-0.54) and the Triadic Primes proof (-0.185), suggesting that participants' judgments of their understandings should not be used as a primary means to assess the efficacy of new proof formats as some researchers have done (e.g., Rowland, 2001).

5. Discussion

5. 1. Summary of main results and caveats

Our findings suggest that the summary provided in Level 1 of a structured proof can be a double-edged sword. On the one hand, if readers are able to understand that summary (which may depend on their familiarity with the subject and the difficulty of the proof) they will have a better chance of grasping the higher level ideas of the proof. Indeed, in the quantitative internet study, we found that participants who read structured proofs performed better on questions pertaining to a proof summary than participants who read an analogous linear proof of the same theorem, suggesting that some students were able to grasp the summary given in Level 1. However, as suggested by our qualitative study, this might not be the case, especially if the reader is not familiar with the ideas in the proof.

Further, we failed to find evidence that structured proofs improve participants' abilities to answer questions pertaining to transfer, justification, or relating the ideas of the proof to specific examples. In fact, while the evidence for each of these three categories was inconclusive, in each case the data suggested an effect in the opposite direction. The results from our qualitative interview study suggest an account for these findings. Students reading a structured proof may have difficulty comprehending the details of the proof, both because mastering the details of the proof requires the participants to integrate information that is far apart in the proof and because some students may have failed to understand the summary of the proof that they had read. These difficulties may have been exacerbated by their lack of familiarity with and understanding of the nature of structured proofs.

Any conclusions reached from this paper should be viewed as tentative for the reasons mentioned when we posed our research questions. We view our data as a contribution to the still open question of whether and how structured proofs improve proof comprehension. While our data are limited, we are not aware of studies that show that structured proofs improve students' appreciation for proof or that delivering structured proofs in a lecture format leads to more learning benefits than presenting these proofs as text. As we argue below, we strongly encourage such research. We hope the results from our qualitative interview study offer guidance on how such studies might be conducted.

5. 2. Significance of results

We discuss how our findings relate to three hypotheses about the pedagogical value of structured proofs: (a) Implementing structured proofs in instruction in a straightforward manner will lead to large and clearly observable learning gains, (b) with careful instruction and preparation, structured proofs have the potential lead to observable learning gains, and (c) structured proofs do not have the potential to lead to significant learning gains.

We believe our data provide evidence against hypothesis (a). Participants who read structured proofs were better able to identify an accurate summary of the proofs than their counterparts who read a linear proof. However, this is most likely because a summary was explicitly provided for them in Level 1 of the structured proof. It is possible that the same benefit could be observed in a linear proof if the linear proof began by providing a summary. We did not observe benefits for other assessment items, meaning that presenting structured proofs as text did not lead to large or observable learning gains. It is possible that using different assessment items or delivering structured proofs as a lecture or in another format would have yielded large and observable learning gains, but as of now, there is no empirical evidence that this is the case.

Some mathematics educators familiar with the challenges of designing instruction might object that it is naïve to expect large learning gains from a relatively simple instructional intervention, arguing that merely tinkering with the format of a proof would not be sufficient to improve student comprehension. However, if we accept this objection, then the suggestions in the literature that structured proofs can improve comprehension should be tempered. We should not recommend that mathematics professors implement structured proofs in their classrooms without also providing them with further recommendations on how they can be implemented effectively. Claims that structured proofs improve comprehension (e.g., Melis, 1994) should be qualified appropriately.

Our data do not distinguish between hypotheses (b) and (c), as we believe a more careful and nuanced implementation of Leron's structured proofs may lead to substantial learning gains. Hence we view it as a worthwhile direction for future research to design and assess instruction based on Leron's structured proofs that demonstrates its effectiveness. We feel that this research is valuable for three reasons. First, if we are going to make research-based suggestions to mathematicians on how to improve their pedagogical practice (as Selden and Selden (2008) do for structured proofs), it is important that we have empirical support for these suggestions. There are instances when theoretically-based pedagogical suggestions lead to minimal, or negative, learning gains, even in the domain of proof presentation (Roy, Alcock, & Inglis, 2010). Second, a positive demonstration of the efficacy of Leron's structured proofs would not only provide evidence *that* this improves learning, but would also provide insight into *what actions a teacher can take to make them work*.

Finally, as Alcock (2010) notes, implementing innovative pedagogy in an advanced mathematics classroom often involves making trade-offs. For instance, it might well be the case that structured proofs lead to increased proof comprehension; however, this only occurs with extended training on how these proofs should be read and with a lengthy, detailed classroom presentation. Ultimately, choosing the time to devote to different activities in advanced mathematics courses involves value judgments by the professor. However, without knowing the types and magnitudes of the learning gains of structured proofs or what a professor would need to do for these to be effective, we are not yet in a position to make these judgments.

REFERENCES

Alcock, L. (2009). e-Proofs: Students experience of online resources to aid understanding of mathematical proofs. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education*. Available for download at: http://sigmaa.maa.org/rume/crume2009/proceedings.html. Last downloaded April 10, 2010.

Alcock, L. (2010). Interactions between teaching and research: Developing pedagogical content knowledge for real analysis. In R. Leikin & R. Zazkis (Eds.) *Learning through teaching mathematics*. Springer: Dordrecht.

Alibert, D. and Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.) *Advanced Mathematical Thinking*. Kluwer: The Netherlands.

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

Cairnes, P. & Gow, J. (2003). A theoretical analysis of hierarchical proofs. In A. Asperti, B. Buchberger & J.H. Davenport (Eds.) *Mathematical knowledge management. Proceedings of MKM 2003*. LNCS 2594. Springer.

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

Dreyfus, T. (1991). Advanced mathematical thinking processes. In D. Tall (Ed.) *Advanced Mathematical Thinking* (pp. 25-41). Dordrecht: Kluwer.

Gosling, S.D., Vazire, S., Srivastava, S., & John, O.P. (2004). Should we trust web-based studies? A comparative analysis of six studies about internet questionnaires. *American Psychologist*, 59, 93-104.

Griffiths, P.A. (2000). Mathematics at the turn of the millennium. *American Mathematical Monthly*, *107*, 1-14.

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

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

Inglis, M. & Mejia-Ramos, J.P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27, 25-50.

Lamport, L. (1995). How to write a proof. *American Mathematical Monthly, 102,* 600-608.

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, 90(3), 174-184.

Leron, U. and Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly*, 102, 227-242.

Mamona-Downs, J. & Downs, M. (2002). Advanced mathematical thinking with a special reference to mathematical structure. In English, L. (Ed.) *Handbook of international research in mathematics education*. Lawrence Erlbaum Associates: Mahwah, NJ.

Mejia-Ramos, J. P. & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna & M. de Villiers (Eds.), *Proceedings of the ICMI Study 19 conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 88-93), Taipei, Taiwan.

Mejia-Ramos, J.P. Weber, K., Fuller, E., Samkoff, A., Search, R., & Rhoads, K. (2010). Modeling the comprehension of proof in undergraduate mathematics. In *Proceedings of the 13th Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC.

Available for download at: http://sigmaa.maa.org/rume/crume2010/Abstracts2010.htm. Last downloaded March 20, 2011.
Mejia-Ramos, J.P. Weber, K., Fuller, E., Samkoff, A., Search, R., & Rhoads, K. (submitted). An assessment model for proof comprehension in undergraduate mathematics.
Melis, E. (1994). How mathematicians prove theorems. In *Proceedings of the 16th Annual Cognitive Science Conference*. Atlanta, GA.
Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. Presentation at the *14th Conference for Research in Undergraduate Mathematics Education.* Abstract downloaded from: http://sigmaa.maa.org/rume/crume2011/Preliminary_Reports.html.
Porteous, K. (1986). Children's appreciation of the significance of proof. *Proceedings of the Tenth International Conference of the Psychology of Mathematics Education* (pp. 392-397). London, England.
Reips, U.-D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 89–117). San Diego: Academic Press.
Rowland, T. (2001). 'Generic proofs in number theory.' In S. Campbell and R. Zazkis (Eds.) *Learning and teaching number theory: Research in cognition and instruction*. (pp. 157-184). Westport, CT: Ablex Publishing.
Roy, S., Alcock, L., & Inglis, M. (2010). Supporting proof comprehension: A comparison study of three forms of comprehension. In *Proceedings of the 13th Conference for Research in Undergraduate Mathematics Education*. Available for download from: http://sigmaa.maa.org/rume/crume2010/Abstracts2010.htm
Schoenfeld, A. (2000). Purposes and methods of research in mathematics education. *Notices of the AMS*, 47*,* 641-649.
Selden A. & Selden J. (2003) Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for research in mathematics education* 34(1) 4-36.
Selden, A. & Selden, J. (2008). Overcoming students' difficulties with learning to understand and construct proofs. In M. P. Carlson and C. Rasmussen (Eds.) *Making the Connection: Research and Teaching in Undergraduate Mathematics Education,* MAA Notes: Washington, D.C.
Thurston, W.P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, 30, 161-177.
Weber, K. (2004). Traditional instruction in advanced mathematics classrooms: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, 23(2), 115-133.
Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12, 306-336.

APPENDIX- Materials used in the quantitative internet study.
Directions for participants:

Recently a mathematician proposed a new way of presenting proofs to make them easier to understand. However this way of presenting proofs has not been tested with students. The purpose of this study is to see how students understand proofs written in this format.

This format presents proofs in levels. Level 1, the top level, gives in very general terms a description of how the proof will proceed. Level 2 carries out the arguments described in Level 1. If there are some logical details or computations for some of the ideas in Level 2, these details may be pushed down until Level 3.

The motivation behind this presentation is that the reader can gain a general sense of what the proof will do before seeing all the detailed arguments.

In reading this type of proof, you may encounter a "between levels" section. In this section, ideas that will be used in later levels are introduced. This way, statements in the proof will be motivated and not appear to come out of nowhere.

Proofs used in the studies:

Structured version

Claim. The equation $x^3 + 5x = 3x^2 + \sin x$ has no nonzero solutions.

Proof:

Level 1. We define $f(x) = x^3 - 3x^2 + 5x - \sin x$. Solutions of $f(x$ precisely correspond to solutions of $x^3 + 5x = 3x^2 + \sin x$. Assume the cl false: then $f(x) = 0$ has a nonzero solution. We show (in Level 2):

a. $f'(x) > 0$ for all x

b. If $f(x) = 0$ has a nonzero solution, then there is a number c for $f'(c) = 0$.

Together, these conclusions clearly produce a contradiction, so the cl proved.

Level 2a. $f'(x) = 3x^2 - 6x + 5 - \cos x$. Using algebra (in Level 3 show this expression is always positive.

Level 2b. Suppose $f(x) = 0$ has a nonzero solution, that is, there $s \neq 0$ and $f(s) = 0$. Since $f(0) = 0$ (Level 3b), this implies there is a that $f'(c) = 0$, contradicting the fact that $f'(x) > 0$. for all x, which w tablished in Level 2a. (The details are given in Level 3c).

Level 3a. We support the claim in Level 2a as follows:

$f'(x) = 3x^2 - 6x + 5 - \cos x = 3(x^2 - 2x + 1) + 2 - \cos x = 3(x-1)^2 + 2 -$

Since $3(x-1)^2 \geq 0$ and $2 - \cos x > 0$ for all real numbers x, $f'(x) > 0$ real numbers x.

Level 3b. $f(0) = 0^3 - 3(0)^2 + 5(0) - \sin 0 = 0$

Level 3c. We support the claim in Level 2b as follows: Rolle's theorem that if a differentiable function f has the property that $f(a) = f(b)$ for then there is a c such that $a < c < b$ and $f'(c) = 0$. In our case, we $f(0) = f(s) = 0$. Hence there is a c between 0 and s such that $f'(c) = 0$.

We define a number as *monadic* if it can be represented as $4j + 1$ for integer j, and *triadic* if it can be represented as $4k + 3$ for some intege triadic (monadic) prime refers to a number that is both triadic (monadi prime. Note that every odd prime is either a monadic prime or a triadic

Claim. There exist infinitely many triadic primes.

Proof:

Level 1. Suppose the theorem is false and let $p_1, p_2, \ldots, p_n$ be all the primes. We construct (in Level 2) a number M having the following prop

a. M as well as all its factors are different from $p_1, p_2, \ldots, p_n$;

b. M has a triadic prime factor.

These two properties clearly produce a contradiction, as we get a triadic which is not one of $p_1, p_2, \ldots, p_n$. Thus, the theorem is proved.

Between levels: How will we define M? It is natural to try $M = p_1 p_2 \cdots$ which meets requirement (a) but not (b). In fact, since M itself may tu to be prime, it must be triadic to meet requirement (b). A natural seconc is $M = 4p_1 p_2 \cdots p_n + 3$. However, since $p_1 = 3$, M is divisible by 3, in vi of (a). This 'bug' is easy to fix: simply eliminate 3 (i.e. p_1) from the p in our definition of M.

Level 2. Let $M = 4p_2 \cdots p_n + 3$. M is clearly triadic. We show tl satisfies the two requirements from Level 1.

Level 2a. Requirement (a) means that no p_i should divide M. I $p_2, p_3, \ldots, p_n$ do not divide M as they leave a remainder of 3, and 3 dc divide M as it does not divide $4p_2 \cdots p_n$.

Level 2b. As for requirement (b), suppose on the contrary that all prime factors were monadic. Then M, as a product of monadic numbers, itself be monadic (Lemma, Level 3), which is a contradiction.

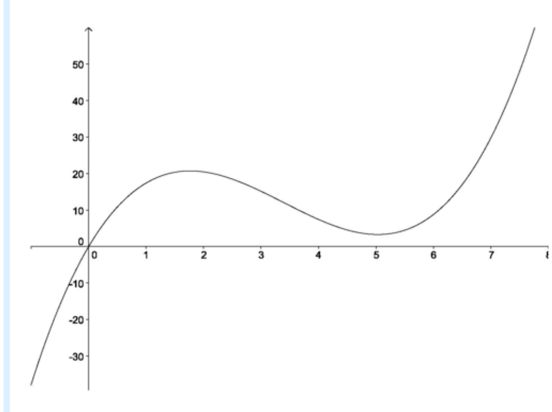
Level 3. *Lemma:* The product of monadic numbers is again a monadic nu *Proof of lemma*: Consider a product of two monadic numbers:

$$(4j+1)(4k+1) = 4k \cdot 4j + 4k + 4j + 1 = 4(4jk + j + k) + 1$$

which is again monadic. Similarly, the product of any number of monadi bers is monadic.

Questions used in the quantitative study:

	Triadic Primes proof	Only Zero proof
Summary	Which of the following is a better summary of this proof, A or B? A. It lists all triadic primes and then uses this finite list of triadic prim obtain a contradiction by constructing a number that is triadic bu no triadic prime factors. B. It lets $M = 4p_2 \cdots p_n + 3$, where $p_i \neq 3$. 2 does not divide M bec M is odd. p_i does not divide M because it leaves a remainder of 3. produces a contradiction. C. I don't know which summary would be better.	Which of the following is a better summary of this proof, A or B? A. $f(x) = x^3 - 3x^2 + 5x - \sin x$. If there were non-zero s for which f then there would be a point where $f'(x) = 0$. However $f'(x)$ is positive. Hence 0 is the only solution. B. Since $f(x) = x^3 - 3x^2 + 5x - \sin x$, then $f'(x) = 3x^2 - 6x + 5 -$ $3(x-1)^2 + 2 - \cos x$. If s is a solution, $f(s) = 0$. A non-zero would mean that there is a c such that $f'(c) = 0$, which is a contra C. I don't know which summary would be better.
Transfer	If one were to adapt the proof to show that there are infinitely many p of the form $6j + 5$, what would be an appropriate definition for M? A first that there are finitely many primes of this form, say $p_1, p_2, \ldots, p_n$ $p_1 < p_2 < \ldots < p_n$. Please type your answer in the space below. For example, if you th that the definition of M should be $2p_1 + 2p_2 + \cdots + 2p_n$, you could e this as "2p1+2p2+...+2pn."	Why is it impossible to use the ideas from the proof to show that the $x = \sin x$ has no nonzero solutions? A. To use Rolle's theorem, the function $f(x) = x - \sin x$ would ne differentiable. B. There is a nonzero solution to this equation. C. $f(x) = x - \sin x$ has critical points. D. None of the above. E. I don't know.
Justification	How was the conclusion that "the product of any number of monadic nur is monadic" used in the proof? A. It was used to show that no triadic prime can divide M. B. It was used to show that M must be monadic. C. None of the above. D. I don't know.	Where was the fact that $f(0) = 0$ used in this proof? I. It was used to show that $f'(x) > 0$ for all real numbers x. II. It was used to show that $f(s) = 0$, where $s \neq 0$, would imply a co tion. A. It was used to show I but not II. B. It was used to show II but not I. C. It was used for both I and II. D. It was used for neither I nor II. E. I don't know.
		How was the fact that $f'(x) > 0$ for all real numbers x used in this pro A. It was used to show that $f(s) = 0$. B. It was used to show that solutions of $f(x) = 0$ precisely corres solutions of $x^3 + 5x = 3x^2 + \sin x$. C. It was not used to show either A or B. D. I don't know.

Application to examples/diagrams	Supposing 3, 7, 11, and 19 were the only triadic primes, which of the foll illustrates the main steps of the proof? ○ A. Let $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3$. 7, 11, and 19 do not divide M since leave a remainder of 3, and 3 does not divide M since it does not $4 \cdot 7 \cdot 11 \cdot 19$. Thus, M has only monadic factors and is monadic. How M is triadic because it is of the form $4k + 3$. This is a contradictio ○ B. Let $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3$. We show that M is divisible by a triadic prime than 3, 7, 11, and 19, which is a contradiction. $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3 =$ The prime factors of M are 5 and 1171. However, $1171 = 4 \cdot 292 +$ 1171 is triadic. This is a contradiction. ○ C. Neither of the above illustrates the main steps of the proof. ○ D. I don't know.	The graph of $g(x)$ is given below. Is it possible that the ideas of the pr be used to show that the equation $g(x) = 0$ has no nonzero solutions?  ○ A. Yes. ○ B. No. ○ C. I don't know.

A Multi-Strand Model for Student Comprehension of the Limit Concept

Gillian Galle
University of New Hampshire
get7@unh.edu

Students demonstrate a variety of behaviors while solving problems involving limits. However, no model currently exists to address what students do understand about limits while accounting for these varied behaviors. This study presents one possible model. This paper proposes that student behavior while solving limit problems may be interpreted by strands that reflect the student's method for solving a problem involving limits, the student's justification for the solution, and the applicability of the student's method and justification within the context of the problem. This paper concludes with the presentation of a model which demonstrates how these strands connect to each other and a proposal of future directions for its refinement.

Key words: limits, student understanding

Introduction

Much research has focused on the misconceptions held by students as they navigate the concept of limits in their first year calculus courses (Davis & Vinner, 1986; Ferrini-Mundy & Graham, 1994; Szydlik, 2000; Tall, 1992; Williams, 2001). However, very few studies seem to identify what students actually do appear to know about limits and the ways in which they demonstrate this knowledge. Hence there is a need for a model that would allow us to determine what students actually know and understand about limits.

Exacerbating the situation further is the condition that, depending on the question, students may exhibit a range of behaviors they make sense of the question and develop a solution strategy (Koedinger & Nathan, 2004; Schoenfeld, 1985). Any attempt to create a model of what students know about limits needs to take this issue into consideration. Despite the inherit difficulty of accommodating the different responses that may be elicited from a student by two differently phrased questions, there are several lenses available in the current literature that give a starting point to the construction of such a model.

One lens addresses the method that a student chooses to utilize in solving a problem involving limits. The Harvard Calculus Consortium's "Rule of 4" (Knill, 2009) aligns with the four main modalities that students exhibit as they work on limit problems.

Another lens looks at the students' justifications of their methods for solution and their subsequent work to produce the solution. Alcock and Weber's (2008) discussion of syntactic and semantic proof strategies lends itself to describing the types of explanations produced by students regarding their solutions.

The combination of these lenses along with the identification of their fit within the context of the problem, or their applicability, provides the basis for a multi-strand model to describe student understanding of limits. This paper provides evidence of such a model's usefulness in identifying student knowledge of limits and concludes with a presentation of the model in its entirety.

Literature Review

Out of the tools available for measuring student understanding of limits, the one cited most often is the 7 Step Genetic Decomposition developed by Cottrill, et. al. (1996). It arose as an aid for observing the progression of levels of student understanding of the formal definition of limit in the realistic mathematics education (RME) environment. The seven stages track the student's grasp of necessary processes and his subsequent coordination of these processes with each other in order to develop a full picture of the formal definition of limit. This model was later expanded by Swinyard (2009) by the addition of more detail to stages 5 – 7 of the original genetic decomposition.

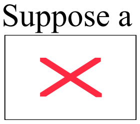
The genetic decomposition was not designed to be a metric or rubric for assessing what knowledge students have about the limit concept. There was no reason to suspect that it would work outside an RME setting, let alone inform us of what a student is thinking. However, it was not unreasonable to believe that there was enough rich description given in the stages to compare to a student's verbalized thought process and possibly assess what the student understood. That is, if students used the language of the formal definition informally as they approached solving problems about limits that did not require the formal definition.

However, the genetic decomposition was not designed to account for the different behaviors that students exhibit when they approach the solution of limit problems that do not involve the formal definition. Hence we need a way to identify and categorize the different student solution methods that arise. One way to do this is to group the exhibited behaviors by their intent through observation of the student's words, symbols, and actions. That is, we can consider the words, symbols, and actions to be external signs of the student's understanding. Thus, in the language of semiotics we can group external signs that indicate similar intent or seem to be the same method into modalities (Saussure, 1966).

The Harvard Calculus Consortium's (HCC) "rule of four" provides a convenient preliminary categorization scheme for these modalities. The HCC was one of the first groups to recognize the importance of being able to approach a problem multiple perspectives and advocate for the inclusion of these perspectives in the classroom. Originally the HCC promoted the "rule of three" and suggested that calculus concepts, such as limits, should be presented graphically, algebraically, and numerically (Knill, 2009). This was later extended to the "rule of four" to include a verbal explanation of the material as well.

If these are the rules that guide explanations of examples in typical calculus textbooks and lectures, it seems reasonable to expect some of these modalities of thought to appear while students approach the solution of problems involving limits. Thus for the creation of the proposed model, the "rule of four" has been adapted from a teaching suggestion into the categories for the method modalities students may exhibit while solving problems.

While the method a student uses to approach a problem may be easy to determine, the explanation he then offers of how he has arrived at the answer can be difficult to decipher. Hence, the model must find someway to interpret the student's justification. One way in which we can look at the student's explanation is to see how it relates to the context of the given problem. Either the student proceeded in a straightforward manner suggested by the context of the problem or the student tapped into knowledge of ideas related to the problem to ascertain an answer prior to returning to the context of the problem. This is the idea developed by Alcock and Weber (2008) when they proposed the idea of semantic versus syntactic proof production.

The syntactic refers to the student staying within the symbolic context given by the problem. Suppose a student was given the problem: "Evaluate the following limit algebraically: ." If the student then chose to manipulate the symbols by factoring the numerator, crossing out like terms, and then equating the statements in order to evaluate the limit, he would be exhibiting syntactic reasoning.

If the student had instead graphed the function to visually ascertain what the limit is doing or if the student had created a table of values to look for a trend in the function's values, he would be using semantic reasoning. Thus, for the purposes of the model, the semantic is when the student moves beyond the symbolic context of the problem to reason about the solution prior to restating the answer in the necessary context.

Method

Using a grounded theory design (Glaser & Strauss, 1967; Strauss & Corbin, 1990), data was collected in two rounds with analysis of the first round of data informing the second round of data collection. Both rounds of data were collected during fall semesters from large lecture, first year, calculus courses. Students enrolled in these courses were typically physics majors, math majors, or engineering majors.

Data Collection - Round 1

A 20 item multiple choice assessment tool was distributed to 55 students. Out of these 20 items, 3 items were determined to be most relevant to the issue at hand. Four students were selected, representing a range of student understanding, to participate in task-based interviews based on their answers to these items.

Task 1.

In this question you are required to encircle the number in front of your choice(s). Which statement(s) in 1.1 to 1.7 below must be true if f is a function for which $\lim_{x\to 2} f(x) = 3$? Encircle 1.8 if you think that none of them is true.

1.1. f is continuous at the point .

1.2. is defined at .

1.3. $f(2) = 3$.

1.4. .

1.5. exists.

1.6. For every positive integer n, there is a real number such that if $0 < |x-2| < \delta$, then $|f(x) - 3| < \frac{1}{n}$.

1.7. For every real number , there is a real number such that if $|f(x) - 3| < \varepsilon$, then $0 < |x-2| < \delta$.

1.8. None of the above-mentioned statements.

Figure 1. Task 1 for the first round of data collection, adapted from Bezuidenhout (2001).

Each interview was approximately 60 minutes in length and followed the same format. Encouraging the use of the talk-aloud protocol, students were presented with two tasks adapted from Bezuidenhout (2001), the first of which is presented in Figure 1, which got at the students' conceptual understandings of limits. The purpose of the tasks was to use conceptual limit problems as a means to explore how well the genetic decomposition described the student's understanding. Of the four students interviewed, one student was unable to complete the tasks. Thus, the results of her interview were excluded from analysis. Of the remaining three interviewed students, two of them were interviewed together. As their insights often played off each other, there may be times when Arlene and Amos[1] are discussed in tandem.

Interviews were transcribed and then coded for evidence of the student's standing within the genetic decomposition. The analysis of this data informed the creation of a 6 item multiple choice assessment tool and a new set of interview tasks for the second round of data collection.

Data Collection - Round 2

During the second round, a 6 item multiple choice assessment, augmented with an "explain your answer" section, was administered to 137 students. Based on student responses to this assessment, 6 students, again representing a range of student understanding, were selected to participate in task-based interviews. Each interview was approximately 60 minutes in length.

The results of the first round of interviews demonstrated that students do draw on a number of modalities for approaching limit problems. Thus the 6 tasks administered during the interviews in this round of data collection were designed to draw out as many of those modalities as possible. The goal was to elicit a wide range of responses from the students by making them think about limits from several different perspectives. These tasks required students to think about everything from "what does it mean for the limit to exist" up through "solve this (unfamiliar) limit problem." Examples of three of the tasks from this round of data collection are given in Figure 2. Students were encouraged to complete these tasks to the best of their abilities using the talk aloud protocol. Interviews were subsequently transcribed and coded for evidence of the student's modality of thinking and the type of justification used.

Task 1: Can you give some examples of a function where $\lim_{x\to -2} f(x)$ exists? Can you give some examples of a function where the limit does not exist?
Task 2: Can you explain why these are (relevant) examples (or non-examples)?
Task 6: (a) How would you evaluate the following limit: ?

Figure 2. Three of the six tasks given to students during the second round of data collection.

Data and Analysis

Although different tasks were used during the different rounds of data collection, both rounds of data offer insight into how students demonstrate their understanding of limits. The students' talk and writing while they worked indicated they were considering a variety of approaches to the problems and often led to more than one attempt at a solution using different modalities. To demonstrate the facility of this model, several students' responses are presented along with the analysis the proposed model allows us to make. The first two students presented come from the first round of data collection and while the remaining two students were participants during the second round.

Amos

Amos primarily exhibited a symbolic modality during his attempts to respond to Task 1. He chose to use the function $f(x) = x$ as an example to help him as he reasoned through which of the statements were true. This led him to some confusion at first. He observed that "So they're saying that as x approaches 2, it's supposed to be equal to 3, but if you replace x with 2, it does not equal 3. So I don't feel that the function's continuous or defined."

Amos's modality was applicable to the problem. The creation of algebraically expressed functions to serve as examples and counterexamples to reason about a given problem shows that Amos understood that it is often easier to discuss limits with a function in hand. His justification during this portion can be classified as semantic as his explanation of his work is centered on his selected example rather than addressing the generalized context presented in the task.

However, his selection of a continuous function that did not fit the context of the problem is a failure of the applicability of his justification. This coupled with his lack of reasoning in the generalized context suggests that Amos may not have a strong conceptual understanding of limits.

Arlene

Arlene started her solution of Task 1 using the verbal modality. That is, as she tried to reason about the veracity of the statements, she spoke her rationalization aloud and left no symbolic record of her thoughts. She said:

> [1.1] should be true because, unless you specify if it's like a right hand or a left hand limit, if there is a limit, I'm pretty sure that the function should be continuous. Otherwise it would have to be like, specified, like an exponential or natural log function or something like that. But I don't think that the second statement, it has to be defined at x equals two, I don't think that one is necessarily has to be. No, that would make the first one wrong.

Recognizing her conflicting statements, Arlene switched to a visual modality to resolve her issues. The graphs she created are presented in Figure 3.

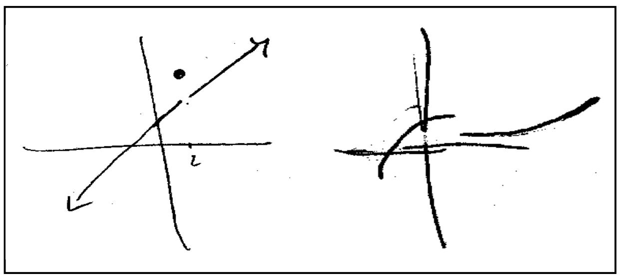


Figure 3. The graphs Arlene drew to reason through Task 1 during Round 1 of data collection.

Arlene explained her new strategy by stating "I was just drawing a graph and I put a dot on it that was supposed to signify three on my little line. I was just thinking, to try and come up with, like an *f* of *x* where it would [...] work for this problem. Like, as *x* approaches two, it would equal to three, just to help me think of some of them."

In Arlene's justifications of her modalities, she stays in the context of the problem by remaining general about the functions she is representing graphically. Although the first graph she drew was a line, her later graph shows that she was also thinking of more generic functions and how they can behave around the value . Thus, her justification is syntactic.

Ultimately, Arlene recognized that both her method and justification were applicable to the problem and used them to rule out statement 1.1. The successful conjunction of all three strands taken together suggests that Arlene has a good understanding of limits in this context. This understanding is further demonstrated through her interactions with her interview partner Amos.

Arlene recognized Amos's selected function as incompatible with the problem. She tried to correct Amos by entering a symbolic modality and utilizing a different equation as she told him to "look at as like any polynomial that could fit this like, I came up with $x+1$." Amos remained unconvinced so Arlene switched to a numeric modality while presenting her graphical examples as more evidence in her favor. She said, "It just doesn't necessarily mean it has to be three at two, because two could be removed. So at like 2.0001 [the function would] still be on that path to approach three."

Barry

Barry also exhibited flexibility in his selection of modalities. When asked to work on Task 1 and 2, Barry started with a verbal description[2]:

> It could have any function that is continuous at *x* equal to -2. It doesn't have to be continuous at that point, but both lines like this have to at least appear to be going towards the same point even when magnified to some extent. The second situation is a little harder to describe but I think it's the same as just checking whether the limit from the left and the limit from the right are equal.

With very little prompting from the interviewer, Barry was also able to produce graphical examples and algebraic expressions for the functions depicted in the graphs he drew. These examples are presented in Figure 4.

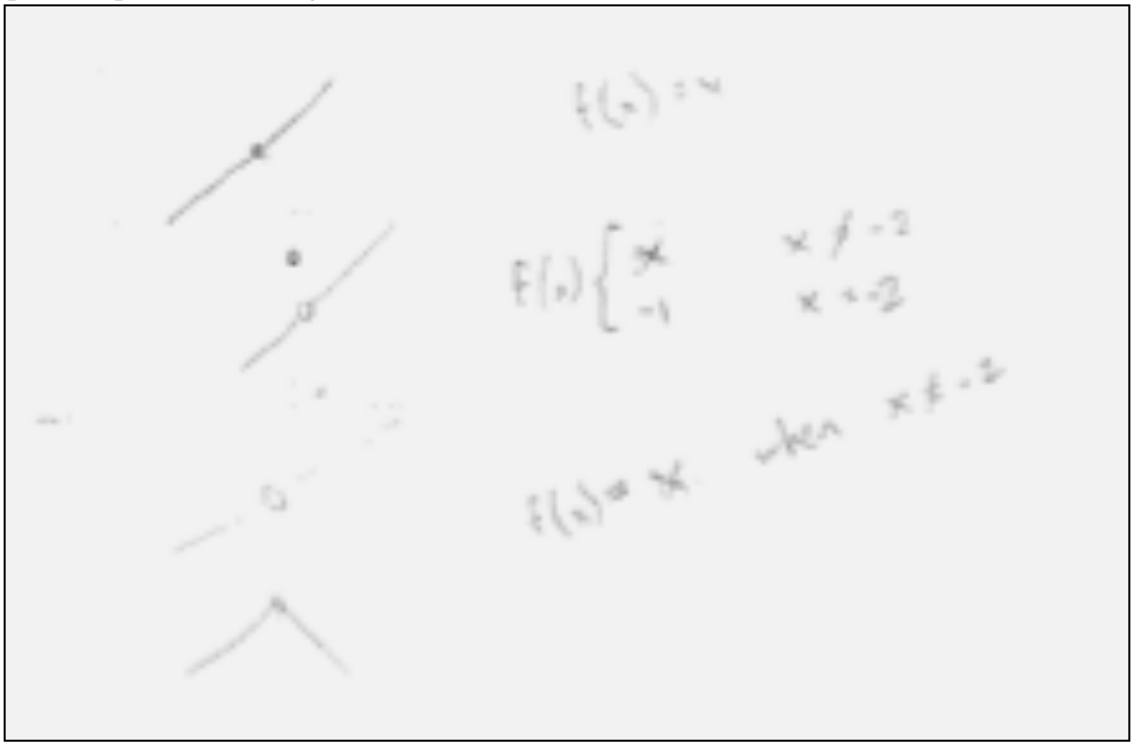


Figure 4. Barry's graphical examples and accompanying algebraic expressions.

While explaining his thought process in the verbal modality and the visual modality, Barry's justification was syntactic. His examples were kept general and represented the families of functions for each case, which was what the context of the task required. When Barry switched to the symbolic modality to create algebraic expressions for the functions depicted by the graph, his justification became semantic with respect to the original problem, as he was no longer being general. However, Barry's justification could still be considered syntactic when taken in the context of the challenge issued by the interviewer, which was to come up with expressions for the functions he drew.

Barry wasn't given much of an opportunity during this task to express whether he recognized the applicability of his justifications beyond the fact that he saw fit to present them in the first place. Thus, Barry exhibited flexibility in selecting applicable modalities to approach solving the problem and followed each modality with an applicable justification. We may be able to conclude that Barry has a strong understanding of the limit concept.

Bud

As he began to respond to the first task Bud also gave a thorough verbal description of when the limit would and would not exist. For instance, Bud said that:

> As long as the point 2 is approached on your function, from both sides of course, then the limit would, as *x* approached 2, would exist […] so basically you can have any sort of shape of line or function or graph or whatever and as long as it approaches 2 from both sides, it'll exist.

Despite prompting by the interviewer, Bud did not feel compelled to switch modalities to re-express his answer. Bud's adherence to the verbal modality in this task may indicate that he was unable to reason about this conceptual question in any of the other modalities. This could be a sign that his conception of limit may have flaws.

During his solution attempts to tasks later in the interview, Bud did demonstrate additional

modalities. Bud started to determine the limit given in Task 6a using a symbolic modality. After determining that the limit would yield an indeterminate form if directly evaluated at the given value, Bud opted to apply l'Hopital's rule. Given that within the classroom norms, the implicit context of this problem would be to apply algebraic methods to its solution, Bud's ensuing justification can be classified as syntactic.

When asked how he would verify his answer, Bud asked if he was allowed access to a graphing calculator. After ascertaining that this would have been a viable option all along, Bud declared, "Alright, well then I would have just started out by graphing [the function]," and proceeded to generate the graph of the function. While waiting for the graphing calculator to finish graphing the function, Bud appeared to be struck by a thought, "So we want the limit as *x* approaches negative 3, so [the graphing window] should be looking at negative 3. So it'd actually be easier for me to go to my table."

Thus for this task, Bud switched from the symbolic modality to the visual modality and subsequently to the numeric modality. Each modality was applicable to the stated problem. Additionally, his justifications were syntactic with respect to the phrasing of the task and were also applicable. However, since Bud exhibited distinct modalities that differed with respect to the different tasks he may not have established connections between the aspects of limits depicted in the two different problems. This could represent a flaw in how Bud understands the limit concept.

Results

The data collected from both rounds of the study suggest that there are three interconnected strands enable us to see student understanding as perceived through their attempts to solve problems involving limits. These strands are method, justification, and applicability.

Method and Modalities

The interviewed students seemed to fall into one of four main modalities as they answered questions involving limits: (1) Symbolic, (2) Visual, (3) Verbal, and (4) Numeric. Each of these modalities encompasses certain unique behaviors, recorded symbols, and words used by students that allow us to differentiate between them.

The symbolic modality is denoted by student creation of examples featuring specific algebraic representations of functions, such as those offered by Amos as he worked through the veracity of a list of statements or those created by Barry when he was prompted to identify which functions he had graphed. This modality also accounts for student use of specific procedures or rules learned about in class. One example of this is Bud's utilization of l'Hopital's rule when trying to evaluate the limit presented in Task 6a in Figure 2.

The visual modality is seen in student use of visual aids, predominantly graphs, for evaluating limits or thinking about families of functions for which a particular limit exists. Hence, Arlene and Barry's use of generic graphs and Bud's graphical analysis of the limit problem in Task 6a are categorized as instances of the visual modality.

The verbal modality is observed when students talk through a solution without making a symbolic record. The data collected during this study relied primarily on audio recordings of interviews, so we are unable to further classify the verbal modalities exhibited by Arlene, Barry, and Bud at this time. It is speculated that students utilizing the verbal modality could arrive at their answer through the use of an analogy or metaphor, by simply talking through a procedure that would have been identified as visual, symbolic, or numeric had it been recorded on paper, or by incorporating gesticulations or other instantiations of perceptuomotor behaviors. A change in data collection strategies will be necessary to record the subcategory of percepuomotor behaviors.

Finally, the numeric modality is evidenced by the student's preference to create tables of values to examine for trends or to more informally plug in several numbers close to the value of interest in order to develop an approximate answer. Though this modality rarely arose as a first choice modality, it often served the purpose of verifying an answer. This is seen in Arlene's use of values close to 2 when trying to convince Amos of the applicability of her graph and also in Bud's approach to answering Task 6a. One reason for the numeric modality's status of second choice may be that the students feel it is not an approach that would be accepted on a homework or exam.

All of these modalities depict the different perspectives a student can hold about limits. Thus each modality has information to contribute regarding how a student has made sense of the limit concept and the approaches that are useful or worthwhile when solving problems involving limits.

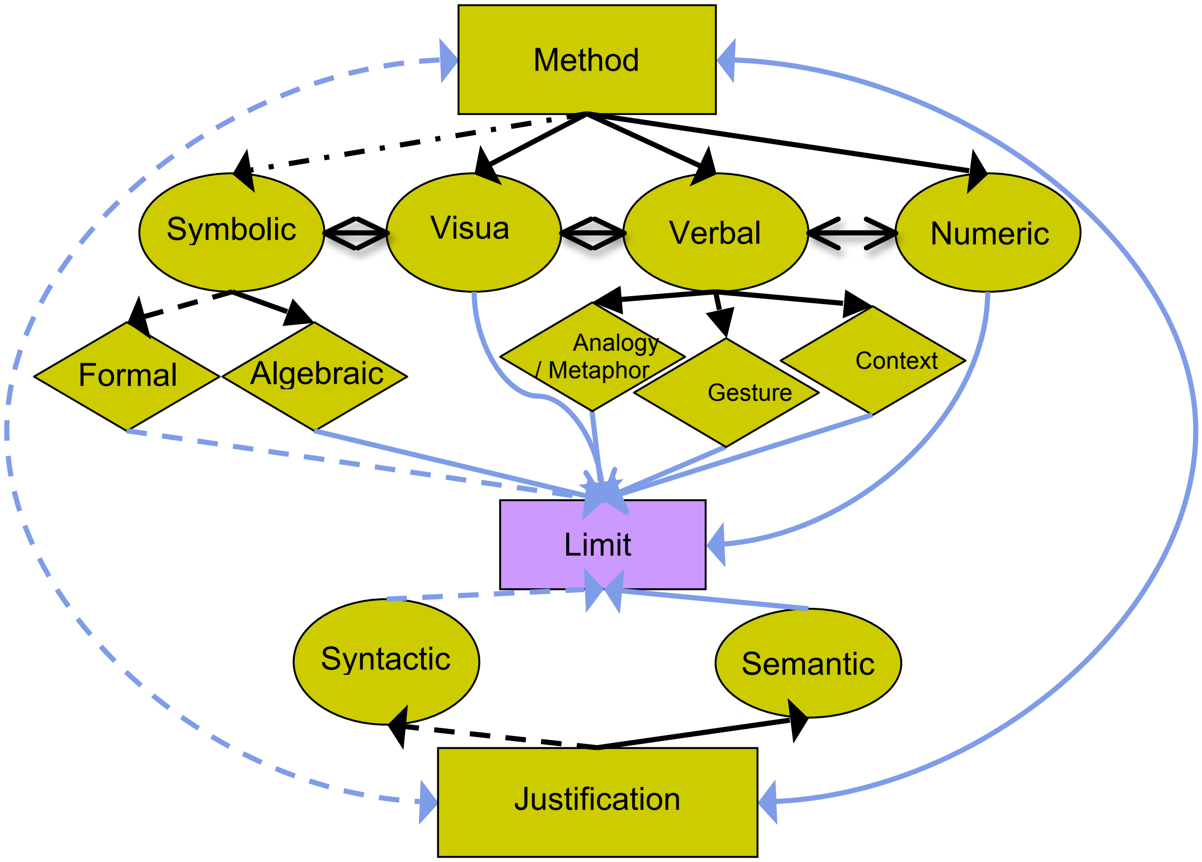
Justification

The method alone is not enough to demonstrate student understanding. Just as important as the student's selection of method is his ability to adequately explain both his work and his selection of method. In particular, justification seems to break down into two main components: whether the student stayed within the framework set up by the question (syntactic) or whether the student translated into another representation to answer the question (semantic).

The students that addressed Task 6a by applying procedures such as multiplying by a conjugate or utilizing l'Hopital's Rule were operating within the implicit framework of the question. Although the phrasing of the task did not explicitly state that an algebraic method must be used, similar limit evaluation questions given on homework and on exams often require such answers. Thus the implicit context was created by the students for themselves. From the students' point of view, the problem was given as an algebraic expression so their work remained algebraic in nature which would be labeled as syntactic justification.

When Bud chose to graph the function and look at a table of values, he left the self-imposed implicit context of the problem and translated into another representation in order to determine the answer to the question. Bud's additional justification during this task could therefore be considered semantic.

Justification also addresses the correctness of the explanation. There is less value to be had in a student that can apply all the algebraic manipulations he was taught but does not realize why the algebraic manipulations allow him to arrive at the correct answer. In a sense, the correctness is the applicability of the justification to the limit at hand.



Applicability

Applicability is the glue that binds a student's method and justification both to each other and to the concept of limit. This strand is important. It is what kept Amos from making progress on his own justifications until Arlene pointed out the error in his thinking. Applicability also arose as students in the second round of data collection debated whether l'Hopital's rule could be used to determine certain limits.

Thus applicability addresses several important issues such as whether the selected method can lead to a solution of the problem or can accurately describe the limit in question, whether the justification validates the method, and whether the student recognizes that his method and justification applies in the context of the problem. If the link of applicability fails to be made along any of these connections, it could indicate that there is a gap in the student's understanding. Then from the instructor's perspective, a failure in applicability should act as an indicator to discuss the solution with the student to try to ascertain where the gap in understanding lies. Even more powerful would be whether the student is able to identify the inapplicability of a method or justification. In this case the gap in knowledge could be more easily accessed. Unfortunately, applicability from the student's perspective can be very difficult to gather data on.

5.4. A Model

It is clear that these three strands are interconnected and relate to each other in an important way. One possible visual representation of the model is presented (below) in Figure 5.

Figure 5. A model of the interconnected strands that comprise student understanding of the limit concept.

In this model, the strands of justification and method are given their own boxes to better visualize their relation to their subcomponents. Applicability then becomes the strand, depicted in blue, which connects the major components of the method strand, the justification strand, and the limit concept. As identified earlier, a lack of connection during a student solution could indicate a flaw in the student's understanding.

The black lines indicate the pieces that are subcomponents of the primary strands of method and justification. Hence justification is either classified as semantic or syntactic. The method strand is then split into the four modalities identified earlier. The further decomposition of the symbolic and verbal modalities merits a brief explanation.

Evidence has been gathered suggesting that the symbolic modality encompasses algebraic behavior. Based on the literature regarding the genetic decomposition (Cottrill, et. al., 1996; Swinyard, 2009) it seems reasonable to stipulate student use of the formal definition could arise given a problem with the correct context. By the very nature of the formal definition, it is most likely to occur in conjunction with the creation of a symbolic record which would necessarily make it a subcomponent of the symbolic modality.

Similarly, due to the limitations of the data collection of this study, the recorded evidence only suggested the existence of the subcomponents of contextual and metaphoric behavior within the verbal modality. Additional data is needed to record physical perceptuomotor behavior exhibited by the student. As such actions would most likely not leave a symbolic record this behavior would fall within the verbal modality.

The dashed lines indicate the connections that are most likely to come into play if a student were to reason through a problem using the formal definition of limit. It falls under the syntactic mode of justification as the language of the formal definition rarely comes into play unless the context of the problem specifically requires it.

Discussion and Implications for Future Research

It seems that different tasks elicit different method modalities for students. If a preference for one modality or a lack of proficiency with a modality were detected, it could show gaps in the student's understanding of the limit concept. At this point it remains a reasonable conjecture that a student that demonstrates greater facility and flexibility when switching between modalities may in fact possess a stronger understanding of the limit concept than a student that is stuck on one modality. Thus the method students select to approach answering a question involving limits needs to be taken into account as it provides one avenue for looking at how the students think about limits in the context of certain problems.

While many first year calculus courses strive for understanding of limits to incorporate the formal definition, it seems students can be successful at determining limits without being able to articulate the formal definition. Hence when exams are designed or we look to formative assessment to inform how we set up our next lecture, we should take the different

modalities students exhibit into account. This may require us to include questions that require the students to not only find the answer, but also to verify their work using another method.

The use of the proposed model to observe student solutions and behaviors while solving problems involving limits offers one way in which to view student understanding. However, it also exposes several other modalities that may need to be looked into. In particular, the modality of perceptuomotor, encompassing gestures under the verbal modality, merits additional investigation to determine what it can tell us about student understanding of limits.

The justification strand also deserves additional exploration and attention. It should be determined whether it can be further disentangled from the method strand. Furthermore, at this time only the subcomponents of semantic and syntactic have been identified. Future research should focus on whether these components are sufficient or if there is more that is encompassed by the justification strand that is not adequately addressed by this model.

Endnotes

1. The names of the students interviewed during the first round of data collection have been changed from the short paper (Galle, 2011). The changes are as follows: Elsa is now Arlene, Jim is now Amos, and Opal is now Agnes.

2. Barry's reference to -2 is not a mistake on his part. The interviewer had incorrectly presented Task 1 as being the limit as *x* approached -2, rather than 2, during this interview.

References

Bezuidenhout, J. (2001). Limits and continuity: Some conceptions of first-year students. *International Journal of Education, Science, and Technology*, *32*(4), 487 – 500.

Cottrill, J., Dubinsky, E., Nichols, D., Schwingedorf, 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*(3), 281 – 303.

Ferrini-Mundy, J., & Graham, K. (1994). Research in calculus learning: Understanding of limits, derivatives, and integrals. In E. Dubinsky & J.J. Kaput (Eds.) *Research issues in undergraduate mathematics learning: Preliminary analyses and results*. Washington D.C.: MAA, 31-45.

Galle, G. (2011). A multi-strand model for student comprehension of the limit concept. In Brown, S. (Ed). Proceedings of the Conference for the SIGMAA on Research in

Undergraduate Mathematics Education, Portland, Or.

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

Knill, O. (2009, February 14). *On the Harvard calculus consortium*. Retrieved from http://www.math.harvard.edu/~knill/pedagogy/harvardcalculus/

Koedinger, K., & Nathan, M. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences*, *13*(2), 129 – 164.

Saussure, F. (1966). *Course in general linguistics*. New York: McGraw-Hill.

Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando: Academic Press.

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

Swinyard, C. (2009, February). Finding a suitable alternative to a potential infinity perspective: A watershed moment in the reinvention of the formal definition of limit. Retrieved on 12/15/2009. Conference on Research in Undergraduate Mathematics Education, Raleigh, NC.

Szydlik, J. (2000). Mathematical beliefs and conceptual understanding of the limit of a function. *Journal for Research in Mathematics Education*, *31*(3), 258 – 276.

Tall, D.O. (1992). The transition to advanced mathematical thinking: Functions, limits, infinity, and proof. In D.A. Grouws (Ed.) *Handbook of research on mathematics teaching and* *learning*. New York: Macmillan, 495-511.

Sociomathematical Norms: Under Whose Authority

Hope Gerson and Elizabeth Bateman
Brigham Young University
hope@mathed.byu.edu

The study of sociomathematical norms initiated by Yackel and Cobb (1996) has become a popular way to make sense of the complexity of mathematical activity in the classroom. In this study we explore the role authority plays in the negotiation and legitimization of sociomathematical norms. We found that sociomathematical norms in this setting were introduced with hierarchal authorities whether the expectation was introduced implicitly or explicitly. We also found, when sociomathematical norms were introduced by a student, mathematical authority played a strong role in the negotiation process including interpretation, and legitimization.

Key words: sociomathematical norms, authority, inquiry-based, calculus, undergraduate

Introduction

In 1996, Yackel and Cobb introduced the study of sociomathematical norms in an attempt to understand how students' mathematical autonomy might be fostered by their mathematical beliefs and values and to make sense of the complexity of mathematical activity in the classroom. They defined sociomathematical norms to be "normative aspects of mathematics discussion specific to students' mathematical activity" (p. 461). In other words sociomathematical norms can be seen as the reoccurring mathematical aspects of discourse that focus on mathematical thinking rather than thinking about mathematics. For example, the social norm that you should justify your answer does not, by itself, insure that your justifications will be accurate, rigorous, or convincing. However, a sociomathematical norm that defines what constitutes a convincing argument can be introduced to set an expectation in the classroom that encourages strong mathematical activity in the form of justification (Yackel & Cobb, 1996; Kazemi & Stipek, 2001).

In 2005, Levenson, Tirosh, and Tsamir found that in a traditional setting with the teacher as authority over the introduction of sociomathematical norms, the teacher and students did not share the same interpretations of a sociomathematical norm. So while, an expectation was introduced, it did not become normative. We suspect that there was a breakdown in the negotiation processs wherein students and the teacher interpret and agree upon the sociomathematical norm. In this study we address the problem identified by Levenson et al. by exploring the role authority plays in the negotiation and legitimization of teacher- and student-initiated sociomathematical norms.

Literature Review

We first discuss research on sociomathematical norms. We distill the literature into a framework that we use to identify and study sociomathematical norms. Finally we share the results of a study in which we analyzed the same data presented here to gain insights into students and teachers use of authority in this inquiry-based setting.

Sociomathematical Norms

Research on sociomathematical norms has covered a sampling of educational settings with most focusing on elementary school mathematics classrooms. Sociomathematical norms have been documented by researchers in first grade (McClain & Cobb, 2001), second grade (Yackel and Cobb, 1996) third grade (Mottier Lopez & Allal, 2007), fourth and fifth grade (Kazemi & Stipek, 2001, Levenson, Tirosh, and Tsamir), junior high (Hershkowitz & Schwarz, 1999), and university (Yackel, Rasmussen, & King, 2000). Some of these classes have been enriched by technology, some have been taught by teachers attempting reformed-minded practices for the first time, and some have been taught by researchers themselves. In all cases, the studies reflect the admonition put forth originally by Yackel and Cobb (1996), "[they] limit [their] discussion to classrooms that follow an inquiry tradition. Nevertheless, sociomathematical norms … are established in all classrooms regardless of instructional tradition" (p 462).

Sociomathematical norms are reported as general categories among which there are multiple ways of defining a norm. For instance, the negotiation of what constitutes an acceptable mathematical explanation (Levenson, Tirosh, and Tsamir, 2005, 2009 ; Mottier Lopez and Allal, 2007; Yackel and Cobb, 1996) and what constitutes a mathematical difference (Mottier Lopez and Allal, 2007; Yackel and Cobb, 1996). These categories are not themselves sociomathematical norms, but represent some of the types of sociomathematical norms that might be present. For each category, there is a continuum of different sociomathematical norms that could be enacted to structure the mathematical activity in the classroom. For example, an acceptable mathematical explanation in one classroom may require only a description of the procedure done to achieve the answer and in another classroom may require a mathematical argument. The mathematical activity expected in crafting and presenting an explanation and in listening to an explanation in these two classrooms would be quite different even though both had established sociomathematical norms that guided the negotiation of what constitutes an acceptable mathematical explanation. The students in the classroom with the stronger sociomathematical norm for explanation will be required to engage in a higher level of mathematical activity.

Originally Yackel and Cobb (1996) focused on the introduction of sociomathematical norms by the teacher. Since then, researchers have documented sociomathematical norms introduced by a research team (McClain & Cobb, 2001), teachers (Yackel, Rasmussen, King 2000), or students (Hershkowitz & Schwarz, 1999). Once introduced, sociomathematical norms are negotiated and re-negotiated by various participants in the class. This ongoing negotiation of sociomathematical norms can be initiated by either student or teacher through interactions: teacher-to-student, student-to-student, or students-to-tool (Yackel & Cobb, 1996, Hershkowitz & Schwarz, 1999). After the expectation is communicated and understood, "whenever someone acts in accordance with [that] expectation, they contribute to the ongoing constitution of the expectation as normative in that situation" (Yackel, Rasmussen, King 2000, p 281).

If the sociomathematical norm is pre-planned and introduced by the teacher (as in: Yackel and Cobb, 1996; Levenson, Tirosh, and Tsamir, 2005) there is little need for a framework that could be used to identify a sociomathematical norm within the social interaction of a class discussion. However, in order to study a sociomathematical norm in its developmental stages, in the classroom, Staub (2007) suggests that one must further focus on the interactions between the teachers and the students as did Mottier Lopez and Allal (2007). At this point we address a concern put forth by Mottier Lopez and Allal (2007) that Yackel and Cobb's (1996) "subtle distinction [between social and sociomathematical norms] does not provide a sufficiently clear foundation for empirically grounded interpretations" (2007, pg 254). While our data and questions are similar to those of Mottier Lopez and Allal (2007) Our solution is not to minimize the difference between social and sociomathematical norms because we agree with Kazemi and Stipek (2001) that "the distinction between social and sociomathematical norms is useful for studying how classroom practices move beyond superficial features of reform" (2001, pg 60). To maintain the strength behind sociomathematical norms and to further articulate the distinction between social norms and sociomathematical norms we have systemized the definition of sociomathematical norms thereby creating a framework through which sociomathematical norms can be identified in videodata of an inquiry-based mathematics class.

As noted above, it is generally agreed that three components must be identified for an expectation to qualify as a sociomathematical norm. Though there are three distinct components, we are not suggesting a hierarchal or chronological progression. The three components of a sociomathematical norm are: 1. a mathematical expectation is set forth (Yackel and Cobb, 1996) 2. the expectation is negotiated by the participants (Yackel & Cobb, 1996, Hershkowitz & Schwarz, 1999), and 3. the expectation becomes normative (Yackel & Cobb, 1996; Hershkowitz & Schwarz, 1999). The negotiation process in component 2 can further be broken down into three sub-components. Within the negotiation process a) a mathematical interpretation of the expectation occurs (Yackel & Cobb, 1996), b) the expectation is agreed upon (Yackel & Cobb, 1996; Yackel, Rasmussen, and King, 2000), and c) the expectation is validated as legitimate. This last subcomponent of negotiation of a sociomathematical norm was never directly discussed in the research but was alluded to by Levenson, Tirosh, and Tsamir (2005), when they suggested that one student, Dan, preferred mathematically based explanations throughout the course even as the teacher explicitly presented the sociomathematical norm that practically-based explanations were preferable. In this case, Dan rejected the teacher's expectation because he did not see the practically-based explanations as legitimate. Therefore, we believe that the validation of the expectation is an important part of the negotiation process that is essential for an expectation to be agreed upon and to become normative.

Since, in this paper, we are mainly concerned with the introduction and negotiation of sociomathematical norms, we feel it necessary to define two new terms. We will refer to a sociomathematical event[4] as an event where a component of a sociomathematical norm is present. So for instance if an expecation is being introduced, that would be called a sociomathematical event. We will define a quasi sociomathematical norm as a group of sociomathematical events that have some but not all of the components of a sociomathematical norm. For example, if an expectation has been introduced and the participants in the class are in the process of negotiating the expectation, but the expectation has not yet become normative, we would call it a quasi sociomathematical norm. This allows us to focus on the developmental stages of a sociomathematical norm without following it to the point at which it becomes normative.

Authority

In previous work, we suggested authority in the classroom hinges on three major concepts: authority relation, legitimacy, and change (Gerson & Bateman, 2010). The authority relation is a

[4] We are using sociomathematical event in essentially the same way as Levenson, Tirosh, and Tsamir (2005) use the term meta-discursive rules. However, we feel that the term sociomathematical event better captures the connection between the components and the sociomathematical norm itself.

relationship between two or more people, with at least one person acting as the bearer of authority and at least one person acting as the receiver of authority. The bearer makes a claim; the receiver recognizes the claim as legitimate and is influenced to change his or her behavior or thinking. In traditional classroom settings, authority is usually hierarchal with the teacher acting as bearer of authority and the student acting as receiver of authority (Herbel-Eisenmann, Wagner, & Cortes, 2008). In inquiry-based classroom settings the bearer and receiver are more fluid roles taken on by both teacher and student at different times (Hamm & Perry, 2002). We define legitimacy of authority as "the knowledge, skills, position, or experiences that influence a person or group within an authority relationship" (Gerson & Bateman, 2010).

In the same work, we identified four different types of authority held by students and instructors

1) hierarchal authority where legitimization is based upon a person's position in the class,
2) mathematical authority where legitimization is based in mathematics,
3) expertise authority where legitimization is based upon the expertise of the bearer, and
4) performative authority where legitimization is based upon the bearer's ability to engage the class. (p. 200)

Of these four types of authority, hierarchal and mathematical authorities both played a major role in the introduction and negotiation of sociomathematical norms as we will further discuss in the results section of this paper. Therefore, we further explain these authorities below. For more information on expertise and perfomative authorities, we refer the reader to our previous paper (Gerson & Bateman, 2010).

Hierarchal authority can be held by students or teachers but has at its source the institutional position of the instructor as bestowed by the institution. A teacher automatically holds this authority and may grant hierarchal authority to a student by inviting her to present to the class. Hierarchal authority is legitimized by the position as instructor or presenter.

We identified two types of hierarchal authority which we called *institutional authority* and *granted authority*. Institutional authority is held by the instructor whose authority is legitimized by her position as instructor of the course and is granted by the institution which hired and assigned her to teach the course. This is a hierarchal authority because the instructor always holds this authority and the students are always subject to it. However, students can also hold hierarchal authority if it is granted to them by someone already holding hierarchal authority. We called this granted authority. For example, when an instructor calls a student to the board to present his solution, the student is temporarily given hierarchal authority over the other students in the class. He has the authority to direct the class discussion and to call on other students to ask or answer questions. Hierarchal authority has nothing to do with the mathematics that is being presented or discussed, but is legitimized by the student's position as presenter which was granted to him by the instructor.

Mathematical authority is invoked whenever a student or teacher explains or justifies a mathematical statement or uses previously explained mathematics to legitimize a claim. The mathematical authority relation is sealed when a student accepts the mathematical argument and changes his thinking or actions to accommodate the argument.

In addition to the two types of hierarchal authority, we identified two types of mathematical authority, *justification authority* and *mathematics community authority*. In both of these cases a mathematical argument legitimizes the claim. When a person holds justification authority, the bearer presents a mathematical argument and the receiver recognizes the legitimacy of the claim through the mathematical argument. In mathematics community authority the bearer invokes previously argued and accepted mathematics to support his claim. The difference between these two authorities is essentially when the mathematical argument is presented.

One key finding of our study on authority was that the receiver is the one who determines if the bearer holds authority, even in hierarchal authority relations (Gerson & Bateman, 2010). The authority relation is sealed when the receiver recognizes the legitimacy of the bearer's claim and is influenced to change. Therefore, even if the bearer has no intention to bear authority, the receiver can bestow authority upon the bearer. And if the bearer acts with intention to bear authority, but is not legitimized by the receiver, there is no authority relation.

Consequently, the classroom teacher may not be able to limit the power of her own institutional authority. For the teacher, all types of authority are naturally subsumed by the institutional authority that she also holds. It is difficult for the teacher to share authority with her students unless she is somehow able to either minimize the influence of her institutional authority or maximize the influence of mathematical authority.

Setting

Our research is set in a teaching experiment (Steffe and Thompson, 2000) in a university honors calculus class conducted by Janet Walter and Hope Gerson. A guiding principle of the design of this class was that agency is at the heart of learning (Brown, 2005; Kohn, 1998; Rogers, 1969; Walter & Gerson, 2007). Agency is defined as "the requirement, responsibility and freedom to choose based on prior experiences and imagination, with concern not only for one's own understandings of mathematics, but with mindful awareness of the impact one's actions and choices may have on others" (Walter and Gerson, 2007). We therefore designed the calculus class to maximize students' exercise of agency. Students worked on tasks designed or selected to elicit conceptually important calculus content without prior instruction. The teaching experiment consisted of two Calculus I classes taught in the fall 2006 and the winter 2007, followed by a Calculus II class taught in the fall 2007.

The corpus of data from this study is taken from two two-hour class periods in the Calculus II class taught in the fall of 2007. In these class periods, students from the four collaborative groups worked on the Hemisphere Tank Task and presented solutions to the class. These class periods were chosen for two reasons. First they occurred early in the semester, in the fourth week of class, as sociomathematical norms were still being negotiated. Second, a compelling episode occurred at the beginning of the second day of the task, where Michael, a student in the class, introduced a new way of thinking about what constitutes a mathematical difference and introduced a new expectation to the class about how mathematical difference should be explored. We recognized this episode as pertaining to the negotiation of sociomathematical norms and wanted to further understand the dynamics in play pertaining to sociomathematical events and quasi sociomathematical norms in this episode. In order to understand this single episode, we felt it necessary to analyze the surrounding mathematical activity with the following questions in mind: 1) what quasi sociomathematical norms are enacted in the two class periods? and 2) under what authority are expecatations introduced? 3) What roles does authority play in quasi sociomathematical norms?

The Hemisphere Water Tank Task

The Hemisphere Water Tank Task (Figure 1) is a common homework problem in Calculus II. It usually appears in a text after the development of the disk and/or shell methods. We used this task quite differently. We gave the task to students before they had learned the disk or shell methods. The completion of the task required students to develop a method for calculating the volume of a solid of revolution and justify their methods and answers.

Hemisphere Water Tank
A certain water tank is a hemispherical bowl of radius 5 feet. Water flows into the tank at the rate of 10 cubic feet per minute.
A. Find the volume of the water in the tank when the water is 3 feet high at the center of the tank.
B. Find an equation for the volume of the water in the tank for any given height, *h*, of the water at the center of the tank.
Extension: How fast is the water level rising when the water is 3 feet high at the center of the tank?

Figure 1: The Hemisphere Water Tank Task

Method

We analyzed four hours of videotape gathered on January 29, 2007 and January 31, 2007. All episodes were transcribed and independently verified by the research team and we proceeded to code the videodata for key ideas, such as authority, agency, social norms, and sociomathematical norms. Together we coded parts of the compelling episode mentioned above. Then we independently coded surrounding episodes. We built consensus about how to define each code and the importance of each code. After identifying sociomathematical events, the first author viewed the 3 weeks of videodata prior to the corpus of data to identify the first occurrence of each quasi sociomathematical norm.

We then used axial coding using the report and key word map features of Transana (Woods and Fassnacht, 2007). During this process we looked for connections among the different key ideas. When connections seemingly emerged from the axial coding, we went back to the original video and transcripts to validate these connections.

Results

We identified sociomathematical events for three quasi sociomathematical norms and one sociomathematical norm during the presentations that defined *what constitutes a good explanation* and *what constitutes a mathematical difference*. In particular, 1) a good explanation consists of connecting the model, context, and equation, 2) a good explanation provides a mathematical argument 2) a different model creates different mathematics, and 3) two equations are different if they act differently on a common model.

What constitutes a good explanation?

In the first presentation, Group One divided the sphere into "a bunch of little cylinders." They found an equation for the volume of each cylinder, $\square=\square\square 2h$, and created a limit of a sum that would calculate the volume of all of the cylinders $\lim \Delta h \to 0 1 h \Delta h \square\square 2 \Delta h$. In his explanation, Paul invoked the quasi sociomathematical norm that a good explanation connects the model, context, and equation. Paul explained his groups' sum, by connecting the model, "the cylinders" in lines 1 and 5, the context "the height of the water" in lines 1 and 5 and the equation

"□□*2*times the height" in line 1. Paul consistently demonstrated that he defined a good explanation as including connections among the model, context and equation.

1 Paul: So to get it so this equation would work, we have to get the height of the cylinders going to zero. It's going from the first cylinder to the height of the water over Δh, where—which would give us the amount of cylinders we have. And times the, uh or no, summing up each little cylinder, just □□*2* times the height of each cylinder, and the limit of a sum. Do you guys get it so far? Questions? [Timbre raises her hand]

2 Derrick: Yeah.

3 Paul: 'Kay, yeah.

4 Timbre: Can you go over the, the *h* divided by Δh? What was that for again?

5 Paul: 'Kay, so the height of the water is right here—is how high each little cylinder is. So if you divide the height by how high each cylinder is, that will give you how many cylinders we have. So you go from the first cylinder to the last cylinder. Do you get that?

6 Timbre: Yeah.

In our next example, the second group was presenting and John asked them to affirm his own interpretations of the variable □*0*. Michael, in line 7, offered an explanation that connected the model and the equation, but did not connect the context. The instructor, Janet, in line 9, implicitly reintroduced the expectation that a strong explanation includes connections among the model, context, and equation.

7 Michael: [□*0*] is the radius of like the cylinders that we are gonna be making here in a minute.

8 John: Okay.

9 Janet: So I also think of that new *r* as if I am looking at the surface of the water at a particular height? It's the radius of that [Michael: of that circle that it makes] circle that the—of the surface of the water.

10 Michael: Right, you take the bowl and you look at it from above, that's the radius of—of the water level at that point.

After hearing Michael's explanation John's response of "Okay" and his subsequent silence indicate that John deemed Michael's explanation as sufficient (or was unable to provide another response that would improve Michael's explanation). Even after the clarification was given and apparently accepted, Janet implicitly suggested a further expectation to connect the explanation of the meaning of □*0* not only with the model of the problem, but also with the context in line 9.

When Paul presented the explanation of his group's equation, lines 1 and 5, he invoked the quasi sociomathematical norm that a good explanation connects the model, context and equation. Later, when Michael presented, his group's equation, he failed to present an acceptable explanation. Janet extended his explanation to include the context, further negotiating and legitimizing (with her institutional authority) the quasi sociomathematical norm of what constitutes a good explanation. This shows that at this point in the class, while there was some agreement on what constitutes a good explanation, the quasi sociomathematical norm had not yet been completely agreed upon and adopted by all of the students in the class. It was still in the developmental stages.

On the other hand, when class members were explaining new mathematical procedures or models, they nearly always included mathematical arguments to support those procedures. For instance, when Paul initially introduced the idea of adding up the volume of a bunch of little cylinders he explained why the limit of the sum, as the height of the cylinders goes to zero, would be necessary.

11 Paul: 'Cause first of all, in our sphere [makes bowl shape in air] the cylinders [uses thumb and finger to show height] wouldn't be exactly cylinders [tilts hands and pantomines rubbing up and down side of bowl] anyways, because the sides would be angled in. [points to picture of bowl cut in cylinders] So to get it, so this equation would work we have to get the height [uses finger and thumb to show height] of the cylinders going to zero.

In this case, Paul essentially argued that in order for the volume of a cylinder to be an accurate approximation of a cross-sectional slice of a bowl, one would have to take the limit as the height of the cylinder goes to zero.

In the next transcript, in line 12, Tyler explained Group Two's model for the first time. Since the model was different than the first model presented, the group was very careful to explain how they constructed their equations. Tyler presented a mathematical argument for $\square2+\square2=25$. Then later in line 14, Heber supplied a mathematical argument for $\square=h-5$.

12 Tyler: We found out that when you use uh, Pythgorean coordinates, you know by saying that this length here [traces the diagonal radius] on this line is always going to be 5, cause from the center point of the sphere pointing to all points on the outside of the sphere, the radius of the sphere is 5. It's always going to be 5. But we are going to have x and y and y could be our $5-h$ or even you could go backwards and do $h-5$. But then we would be able to find that and just using you know $\square2+\square2=25$. Because you have that, you can find out what x and y are.

13 Michael: That's—is everybody clear on the, we just redefined the same variable names that had been used—looks like before, but we defined them as different things. And to understand everything that we are going to do, we need to understand what those represent. So, [Heber: Yeah.] That's kind of important.

14 Heber: So is it clear how we got y equal to h minus 5?

15 Daniel: Um, can you just re-go over that, like write out the steps that you did

16 Heber: We know that $\square=h-5$ 'cause, uh, like if it was negative 5 the height would be 0. So zero minus five is negative five and that's where it is on the y-axis [pointing to diagram]. Or also like the example here, y would be -2 like they had done on their's. Uh, so if y is -2 the height is 3, $3-5=-2$. So just did this to convert y into h. And so we were able to substitute this $h-5$ value into our equation up here [points to equation].

Because the students were regularly providing a mathematical argument to support new models and procedures, we have evidence that in the first three weeks of class, the expectation that an explanation of a new mathematical procedure must include a supporting mathematical argument, had already become normative. Therefore we can say that this was an instance of a sociomathematical norm for what constitutes an acceptable explanation.

What constitutes a mathematical difference?

At the beginning of the second presentation, Tyler and Heber explained their model. As the other group did, they used *r* to represent the radius of each little cylinder. And in contrast to Group One, who used *h* and *y* interchangeably, group two used *y* to represent the *y*-coordinate and *h* to represent the height of the water. Michael suggested, in line 13, that "we defined them [the variables *x* and *y*] as different things." This invoked the expectation that different models create different mathematics.

After Michael invoked the quasi sociomathematical norm that a different model creates different mathematics, Heber, in line 16, continued to explain his group's connection between *y* and *h*. Despite Daniel's suggestion that he "just…write out the steps that you did" Heber's explanation connected the model, context, and equation, further normalizing the quasi sociomathematical norm of what constitutes a good explanation.

Group One solved the problem using the equation, $-5-2\square(25-h2)\square h$, and group Two solved the problem using the equation, $\square 25-h-52\square h$ .After the two groups presented their solutions, Michael introduced a new expectation (lines 17 and 20) for what constitutes a mathematical difference. Heber and Tyler interpreted the expectation (lines 18 and 19) and Michael began to negotiate the meaning of the new expectation (line 20).

17 Michael: Now, we it looks like we've got two different equations? *{expectation}*
18 Heber: Yeah. *{interpretation}*
19 Tyler: But they're the same. *{interpretation}*
20 Michael: Um, they're not the same equation, they both model something different, and I'm 90 percent sure that I know what that is, the difference. *{expectation and negotiation}*

In order for this new expectation that two equations are different if they "model something different" to become normative, Michael's expectation needed to be interpreted, agreed upon, and legitimized. For the next ten minutes, Michael and the class continued to negotiate the meaning of Michael's expectation of what constitutes a mathematical difference. For example in the next excerpt, Michael re-states the expectation embedded in a mathematical argument in line 23, and Robert interprets that to mean solving part A, and expresses that he understands Michael's expectation.

21 Michael: if you just get the volume function and just start evaluating the volume function *{mathematical argument}*
22 Derrick: [inaudible] you could get a general equation *{interpretation}*
23 Michael: Right, but I'm saying take an indefinite in, er a definite integral of their equation. What would that model? What, if you plug in the value of one, into their indefinite integral [sic], what does that represent? *{mathematical argument, mathematical authority, t and restatement of the expectation}*
24 Robert: 'Cause, 'cause do you want me to solve part A? 'cause it [inaudible] [moves thumb and index finger together] I see what you're asking. *{interpretation}*

Michael's initial statement in line 21, that "they both model something different" did not supply enough mathematical information for Robert and the rest of the class to build a mathematical interpretation of the expectation, nor to judge whether it was legitimate. When Michael began to explain his expectation he made explicit how he wanted to compare the two equations, and offered legitimization of the expectation through mathematical authority. Although Michael held granted authority to present his ideas to the class, and expertise authority for his own solution, it was Michael's mathematical authority that legitimized his expectation through his mathematical argument, and allowed others to begin to interpret and agree on the expectation.

In the next excerpt, Michael re-stated the expectation in line 25, and Michael and Heber provided an explanation that included connections among the model, context, and equation as well as a mathematical argument for how they were interpreting the two equations. They explained that they used their own procedure on both equations to get the same answer.

25 Michael: They're both giving the right answer, but they're different equations.
26 Heber: Yeah, we even eval [clears throat], we even evaluated at a different [clears throat] at a different height, for the tank. We evaluated it at one. We took their information at, er we just did another definite integral with theirs, and we went from negative five to negative four with um
27 Michael: Um hmm. [Janet Um hmm] Which is from there [points to the bottom of the tank] to there [points to the first horizontal line up].
28 Heber: We'll call that □□□□. [Justin moans] We did that with theirs and we did the same thing with ours, we went from zero to one, um □□□□ and we got the same exact answer for that so their equation models same, the same exact volume that ours does, uh, which doesn't completely make sense to me, but I'm sure someone understands it.
29 Michael: Alright. So that that's what I'm trying to resolve. Now, what I, what I wanna ss, uhh, think about is if we take an indefinite integral, of the equation that they had, no? That's what we did here, we took an indefinite integral and without an interval we simply got volume as a function of *h*. So if we do the same thing with their equation and take volume as a function of *h* Now, let's analyze, if we plug in an *h* value of one into our equation it will return this volume down there, is that making any sense?
30 Timbre: It seems like the same equation.
31 Heber: Oh, I get what you're saying.
32 John: I'm [sighs]

And while Heber expresses confusion in line 28, he finally indicates that he understands what Michael means by mathematical difference in line 31. Nevertheless, Timbre, in line 30, and probably Justin and John, in lines 28 and 32, as evidenced by their moan and frustrated sigh, have still not been able to interpret Michael's meaning for mathematical difference. Therefore, what constitutes a mathematical difference is still being negotiated and this qualifies as a quasi sociomathematical norm.

Here we point out, that Michael's quasi sociomathematical norm is at its budding stage. In the previous examples, the expectations of what constitutes a good explanation were first introduced in the first week of class. So those expections had gone through a longer negotiation prior to their invocation on this day in class. Michael's expectation, on the other hand, had only been considered for about 10 minutes. Therefore, the lack of agreement among students is not surprising and should be noted before one makes generalizations about the relative strength or weakness of the agreement within the negotiation process

The Role of Authority

As we were studying both the introduction and negotiation of sociomathematical norms, and the authority with which one introduces sociomathematical norms, we noticed that authority was playing two roles. Authority was needed both to introduce and to legitimize a sociomathematical norm.

Michael's initial statement in line 20, that "they both model something different," did not supply enough mathematical information for the class to build a mathematical interpretation of the expectation, nor to judge whether it was legitimate. When Michael began to explain his expectation he made explicit how he wanted to compare the two equations, and offered legitimization of the expectation through mathematical authority. Although Michael held granted authority to present his ideas to the class and used it to introduce the expectation, it was Michael's mathematical authority that legitimized his expectation through his mathematical argument, and allowed Heber to interpret the expectation.

Michael's introduction of the quasi sociomathematical norm of what constitutes a mathematical difference led to over ten minutes of explicit interpretation and negotiation among the participants. We believe that there are three contributing factors to the richness and depth of the interpretation and negotiation process in this case. First, the expectation was explicitly stated. Second, the quasi sociomathematical norm was introduced by a student. Third, the quasi sociomathematical norm was legitimized through mathematical authority.

On the other hand, in the negotiation of the quasi sociomathematical norm that a good explanation connects the model, context and equation, no member of the class ever directly stated the expectation. Instead students and instructors put forth mathematical explanations some which were accepted and some which were questioned or extended by members of the class or the instructors. When the explanations were accepted, they contributed to the agreement of the expectation and further contributed to the expectation becoming normative. However, when the explanations were not accepted, they contributed to the ongoing negotiation process of interpretation and agreement. While mathematical arguments were presented as a part of mathematical explanations (as in lines 1, 5, 11, 12, and 16), these arguments were not used to introduce or legitimize the quasi sociomathematical norm. Instead if any authority was used at all it was hierarchal authority in both its forms.

In fact, we saw very little evidence that this norm was being explicitly negotiated. What we saw, was that each time an explanation was made, either it was accepted implicitly by the class, or someone, often one of the instructors, offered a clarifying remark or question. Over time, with many explanations being offered there was ample opportunity for students to interpret the expectation which likely led to the fairly strong agreement that we witnessed among the participants for what constituted an acceptable explanation. However, the negotiation process was implicit. The lack of both explicit statement of the expectation and the use of hierarchal authority to legitimize the quasi sociomathematical norm could have very easily led to lack of agreement among the participants of the course.

Conclusions

We were able to identify the sociomathematical norm that *a good explanation of a new procedure or model requires a mathematical argument or mathematical authority*. In addition we identified three quasi sociomathematical norms. First, that a good explanation connects the model, context, and equation. Second, that *a different model creates different mathematics*, and third, that *two equations are different if they act differently on a common model*.

We found that both sociomathematical norms and quasi sociomathematical norms in this setting were introduced with hierarchal authorities whether the expectation was introduced implicitly or explicitly. We also found, in the case of the quasi sociomathematical norms that were introduced by Michael, a student in the class, mathematical authority played a strong role in the negotiation process including interpretation, and legitimization. And in fact, in the quasi sociomathematical norm that was introduced implicitly, authority did not play an observable role in the negotiation process and in fact, the negotiation process was implicit. We cannot tell from our data whether the negotiation process would have been more likely to be explicit had the expectation been made explicit or if the expectation had been made by a student. However, we did see a strong and explicit negotiation process in the case where the expectation was made explicit by a student and legitimized with mathematical authority.

In our first study on authority (Gerson & Bateman, 2010), we found that "in a mathematics class, it is those authorities that legitimize claims while also providing convincing evidence that have the highest potential to affect a student's mathematical understanding and mathematical autonomy" (p. 206). We believe that this is also true for the negotiation process of quasi sociomathematical norms. That is, that mathematical authority has the highest potential to allow for rich negotiation leading to strong agreement among the participants.

When instructors introduce sociomathematical norms, as in the study by Levenson, Tirosh, and Tsamir (2005), their hierarchal authority may legitimize the sociomathematical norm, but does not necessarily encourage students to interpret and come to agreement. Therefore students may be more likely to accept the norm before they agree on its mathematical meaning. Therefore, if teachers introduce a sociomathematical norm, they should be aware of the potentially obstructive role their hierarchal authority may play in the negotiation of that norm. The exercise of mathematical authority, through mathematical argumentation, by the instructor and students may influence the extent to which sociomathematical norms are truly agreed upon by both instructor and students.

Hierarchal authority also played a role in negotiating and legitimizing the sociomathematical norm of what constitutes a good explanation. But in this case, the norm was not explicitly negotiated. We believe that there were two likely causes. First, the expectation was never explicitly stated. And second, the hierarchal authority was only legitimized by the bearer's position in the class and therefore did not necessarily carry any mathematical weight through mathematical argument. In addition, we believe that quasi sociomathematical norms are best introduced explicitly so that the negotiation process can also be made explicit. Ideally a sociomathematical norm would be explicitly introduced by a student or instructor, through a hierarchal authority, and then legitimized by a mathematical argument, through mathematical authority. Additionally, if it is introduced by a teacher, we suggest that the teacher should find ways to introduce mathematical authority so that the quasi sociomathematical norm is not accepted without negotiation due to the teacher's institutional authority.

References

Brown, T. (2005). Shifting psychological perspectives on the learning and teaching of mathematics. *For the Learning of Mathematics, 25*(1), 39-45.

Gerson, H. and Bateman, E. (2010) Authority in an agency-centered, inquiry-based university calculus classroom. *Journal of Mathematical Behavior 29*(4), 195-206.

Herbel-Eisenmann, B., Wagner, D. Cortes, V. (2008) Encoding Authority: Persvasive Lexical Bundles in Mathematics Classrooms. Proceedings of the Joint Meeting of the International Group for the Psychology of Mathematics Education and the International Group for the Psychology of Mathematics Education North American Chapter (Merida, Mexico, July 17-21, 2008)

Hamm, J. V., & Perry, M. (2002). Learning mathematics in first-grade classrooms: on whose authority? *Journal of Educational Psychology, 94*(1), 126-137.

Hershkowitz, R., and Schwarz, B. (1999). The emergent perspective in rich learning environments: Some roles of tools and activities in the construction of sociomathematical norms. Educational Studies in Mathematics, 38, 149-166.

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

Kohn, A. (1998). Choices for children: Why and how to let students decide. *In What to Look for in a Classroom.* San Francisco: Jossey-Bass.

Levinson, E., Tirosh, D., and Tsamir, P. (2009) Students' perceived sociomathematical norms: the missing paradigm. The Journal of Mathematical Behavior 28, 171-187.

Levinson, E., Tirosh, D., and Tsamir, P. (2005) mathematically and practically-based explanations: individual preferences and sociomathematical norms. International Journal of Science and Mathematical Educations 4, 319-344

McClain, K. & Cobb, P. (2001). An approach for supporting teachers' learning in social context. In F. Lin, and T. Cooney (Eds.) Making sense of Mathematics Teacher Education (pp. 207-232). Dordrecht, the Netherlands: Kluwer.

Mottier Lopez, L & Allal, L. (2007). Sociomathematical norms and the regulation of problem solving in classroom microcultures. International Journal of Educational Research, 46 252-265.

Rogers, C. R. (1969). *Freedom to Learn: A view of what education might become.* Columbus, OH: E. Merrill Publishing.

Staub, F. (2007). Mathematics classroom cultures: methodological and theoretical issues. International Journal of Educational Research (46) 319-326

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.

Walter, J. & Gerson, H. (2007). Teachers' Personal Agency: Making Sense of Slope through Additive Structures. *Educational Studies in Mathematics, 65*, 203-233.

Woods, D, and Fassnacht, C. (2007). Transana v2.2x. http://www.transana.org. Madison, WI: The Board of Regents of the University of Wisconsin System.Yackel, E. and Cobb, P. (1996) Sociomathematical norms, argumentation, and autonomy in mathematics. Journal for Research in Mathematics Education, 22, 390-408.

Yackel, E., Rasmussen, C, King, K. (2000) Social and sociomathematical norms in an advanced undergraduate mathematics course. Journal of Mathematical Behavior, 19, 275-287.

TRANSITIONING FROM CULTURAL DIVERSITY TO CULTURAL COMPETENCE IN MATHEMATICS INSTRUCTION

Shandy Hauk
WestEd & University of Northern Colorado
shauk@wested.org

Nissa Yestness
University of Northern Colorado
nissa.yestness@unco.edu

Jodie Novak
University of Northern Colorado
jodie.novak@unco.edu

We report on our work to build an applied theory for intercultural competence development for mathematics teaching and learning in secondary and tertiary settings. We use research in social anthropology and communications to investigate the nature of intercultural competence development for mathematics instruction among in-service secondary mathematics teachers and college faculty participating in a university-based mathematics teacher professional development program. We present results from quantitative and qualitative inquiry into the intercultural orientations of individuals and some groups (teachers, teacher-leaders, university faculty and graduate students) and offer details on the development of case stories for use in the professional development of mathematics university teacher educators, in-service teacher leaders, and secondary school teachers.

Keywords: teacher preparation, intercultural competence, professional development, diversity

> *The great challenge for professional learning is that [a learning] experience occurs where design and intention collide with chance.* (Shulman, 1998)

> *I wanted to explain why some people seem to get a lot better at communicating across cultural boundaries while other people didn't improve at all, and I thought that if I were able to explain why this happened, educators could to a better job of preparing people for cross-cultural encounters.*
> (Bennett, 2004)

Lee Shulman (1998), in reviewing the education of professionals (e.g., doctors, lawyers, teachers, clergy) in the 20th Century, noted six defining characteristics of a profession:

1. the obligations of *service* to others, as in a "calling";
2. *understanding* of a scholarly or theoretical kind [e.g., mathematics, pedagogy];
3. a domain of skilled performance or *practice*;
4. the exercise of *judgment* under conditions of unavoidable uncertainty [e.g., instructional decision-making in classroom contexts];
5. the need for *learning from experience* as theory and practice interact; and
6. a *professional community* to monitor quality and aggregate knowledge. (p. 516)

These characteristics play out in a variety of ways throughout engagement in a chosen profession and cultural implications are woven through all six. The first is seen by many as a given (especially in light of teacher pay). The second is addressed in the college degree expectations for pre-service teachers and the continued focus of in-service teacher professional development on learning content, pedagogy, and pedagogical content knowledge. The third and sixth points have had increasing attention in the recent past (e.g., "Mathematical Quality of Instruction" observation protocols, Hill, 2010, and the drive to develop professional learning communities, Borko, 2004). The fourth and fifth items in the list are now emerging as areas for research and foci for teacher professional development. In particular, "human judgment always incorporates both technical and moral elements, negotiating between the general and the specific, as well as between the ideal and the feasible" (Shulman, 1998, p. 519).

Given the diversity of students in the nation's classrooms, teachers in U.S. schools are destined to have opportunities for daily cross-cultural experiences that, for most, will be fraught with unavoidable uncertainty. Now, consider the current state of the art in teacher professional education. What is overtly, clearly, and explicitly offered to pre- and in-practice teachers for being aware of the unavoidable uncertainty in their work, much less about how to make morally and contextually complex judgment calls? Not much. In part this is evidenced by the new teachers who leave the profession within a few years, citing as the reason that they feel they were not prepared for what the work is really like by the teacher education program (Keigher, 2010). If characteristic 4 is not well addressed in teacher education, then an alternative is category 5: Learn it from experience. Learning from experience requires a multifaceted mirror of reflective practice, one that teachers can use to see cross-cultural encounters *as opportunities* to learn.

Background

While the significance of diversity as a factor in the education of American children has been widely discussed for many years, the nature of "diversity" continues to evolve in U.S. classrooms (Aud, Fox, & KewalRamani, 2010). Further, while a similar diversity is evident in some school staffing (e.g., paraeducators, in-class assistants), the teacher and administrator populations continue to be more homogeneous than varied in terms of government-surveyed categories of identification and experience like race, education, and socialization (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006). Many reports from research and practice indicate that culture is a significant factor in the inequities of persistence and achievement in education (e.g., for research see Greer, Nelson-Barber, Powell, & Mukhopadhyay, 2009; practice, Equity Alliance, www.equityallianceatasu.org). From anti-racism training to culturally responsive pedagogies, teacher professional development efforts have emerged largely from the same arena as teacher education itself: psychology. Yet there is another area of the academy from which professional educators can draw great insight: anthropology (Ladson-Billings, 2001). That is, while psychology tackles teacher education through an approach that catalogues and attempts to change a teacher's classroom *disposition* through focused reflection on *behavior*, social anthropology offers the idea of movement along a developmental continuum of *orientation* through focused reflection on *communication* in intercultural experiences. Several frameworks exist for professional contexts that involve understanding, interacting, and communicating with people across various cultures. In particular, healthcare professions and international relations groups have generated suggestions for cultural competence and communication based on theories of intercultural development and conflict resolution styles (e.g., Bennett, 1993, 2004; Hammer, 2005, 2009; Kramsch, 1998; Leininger, 2002; Wolfel, 2008). The core of the orientation-communication approach is building skill at

> *Short definition of culture*: A dynamic social system of values, beliefs, behaviors, and norms for a specific group, organization, or other collectivity; the shared values, beliefs, behaviors, and norms are learned, internalized, and changeable by members of the society (Hammer, 2009).

establishing and maintaining relationships in culturally diverse contexts.

That is, while current teacher education focuses on how an individual teacher can build a classroom community or a professional learning community with certain target characteristics, an orientational framing to teacher education unpacks “community” – how it is defined by teacher(s), staffs and students – and applies attention to the characteristics of the relationships formed among teachers, students, staff, and their respective outside-of-school experiences. The relational considerations of orientation to the world include how we are aware of ourselves and each other, the relationships we perceive, value, and engage, and the various forms of communication we might use to build productive relationships for teaching and learning.

Though some teachers have largely monocultural classrooms, in the sense that most students share experience of a particular set of cultural-general norms and practices, the nature of “diversity” in the U.S. is shifting from such segregated monocultural circumstances to cultural heterogeneity. For example, the 21st century version of multi-cultural can mean 2, 5, even 10 different home language groups in a single classroom (Aud et al., 2010). Cobb and Hodge (2010) explored the development of equitable classroom practices by distinguishing between (a) “cultural alignment” approaches – where a teacher is supported to offer instruction in ways aligned with the homogeneous, “local” culture of students – and where the uncertainty inherent in heterogeneity is under-attended; and (b) a “classroom participation” approach where curricula derived from majority group normative policies (e.g., the *Principles and Standards for School Mathematics*, 2000) drive learning activity and the teacher is expected to provide acculturative support that “might enable particular groups of students to participate substantially in these activities” (p. 13). However, in the “classroom participation” approach, the uncertainty inherent in the heterogeneity of student (and teacher) experience is under-attended. What we offer here attends to the missing aspect of heterogeneity, dealing with the realities of negotiating the multiple cross-cultural relationships in the classroom. Our approach is through an exploration of teacher experience as a foundation for the development of case-stories. In terms of the characteristics of a profession and professional education:

> As a pedagogical device, cases confront novice professionals with highly situated problems that draw together theory and practice in the moral sea of decisions to be made, actions, to be taken. Options are rarely clean; judgments must be rendered. Cases are ways of parsing experience so practitioners can examine and learn from it… and can become the basis for individual professional learning as well as a forum within which communities of professionals can store, exchange, and organize their experience. (Shulman, p. 525).

Motivating Example

To motivate later discussion, we offer the example shown in Figure 1 of a classroom interaction between two students and a professor in a discrete mathematics class. Patricia and Mark are working together to solve a graph theory prompt. Dr. Denton is walking around the room answering student questions and checking in with groups. Figure 1 gives the utterances of Dr. Denton, Patricia, and Mark along with the associated actions performed by Mark and Patricia. The example is fictionalized from teacher self-reports. We will later discuss this example of mathematical interaction using the intercultural competence framework presented in the Conceptual Framework section. Implicit in the vignette are relational comparisons among ideas. Notice that Dr. Denton is comparing what students are doing to a preferred answer in his head. Patricia is noticing the ways Mark's answers are different from hers and from what Dr. Denton says. Mark is looking for how things are the same. Dr. Denton is relating student expressions to things in his head. Patricia is relating actions and comments to what is correct. Mark is interrelating his own work, Patricia's work, and Dr. Denton's statements and noticing in what ways they are all the same. This is a very basic example of three of the five stages in the conceptual framework of intercultural competence.

Setting: Discrete Mathematics Class.

Problem: How many edges are there in a planar connected graph with 5 vertices and 4 faces? Draw such a graph.

Dr. Denton - Discrete Mathematics Professor
Patricia - Pre-service Secondary Math major
Mark - Pre-service Elementary Ed major

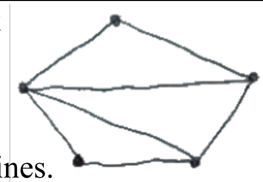
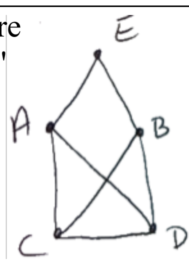
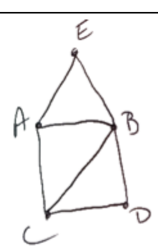
Description of actions while working on the prompt.	*Utterances*
Mark starts drawing vertices and connecting them.	**Mark: Let's see if this even makes sense. Can we draw a picture of it?**
Patricia writes Euler's Formula: (If G is a connected planar graph with V vertices, E edges, and F faces then $V - E + F = 2$), fills in values and solves.	**Patricia: Let's just use the equation like we are supposed to.**
Mark and Patricia look at the prompt again. Mark looks at his drawing of a pentagon and adds two interior lines. 	**Patricia: Okay, I guess we do need to draw a graph.**
Patricia makes a drawing with 5 vertices and connects them with 7 edges. 	**Dr. Denton: If you put the four dots in a square with a dot on top like a house with a roof, it will be easy to grade.**
Mark draws a new picture that looks like a "house," but two of the edges intersect. 	**Mark: Do you mean like this?**
Patricia and Mark look at each other's drawings while Dr. Denton is answering Mark's question.	**Dr. Denton: No. You have to have a roof.**
Mark erases one line and instead connects different vertices 	**Patricia: Yours has 5 faces, edges aren't supposed to intersect. You have to erase one of those in the middle and make a roof line.**
	Dr. Denton: Yes, Mark's drawing is right.

Figure 1. Sample interaction from a Discrete Mathematics class.

Conceptual Framework

The working definition of "culture" (box on page 1), can include professional and classroom environments as well as personal or home experience. Our work to build an applied theory for intercultural competence development for mathematics teaching and learning in secondary and tertiary settings is based on the *Developmental Model of Intercultural Sensitivity* (Bennett & Bennett, 2004). As a developmental model, it ranges from monocultural to intercultural orientations with descriptions of the transitions among intermediate orientations. Figure 2 gives the five orientations and images that we use below as visual metaphors in describing the model.

Figure 2a.

Figure 2b.

Figure 2c.

Denial ⇨ Polarization ⇨ Minimization ⇨ Acceptance ⇨ Adaptation

Figure 2. Stages and visual metaphors for the intercultural development continuum.

The left endpoint of the developmental continuum of orientations is a lens for perceiving the world based in the assumption "Everybody is like me." Though called "denial" by Hammer et al., the orientation might more appropriately be called "innocent" or "bemused." A person with this orientation to culture may become aware of observable differences (e.g., distinctions in food or dress) but not notice more complex difference (e.g., in values, beliefs, or communication norms) and will avoid or express disinterest in cultural difference. A hint of this can be seen in Dr. Denton's focus on a particular representation as the referent in relating to students' efforts.

The transition to the next orientation comes with the recognition of self as distinct from "other" through a noticing of difference, as in awareness of light and dark in viewing a situation (e.g., Figure 2a). The "polarization" orientation is driven by the assimilative assumption "Everybody *should* be like me/my group" and is an orientation that views cultural differences in terms of "us" and "them." Polarization can take the form of "defense" or "reversal." Defense includes a sense of belonging to a group along with an uncritical view towards the values and practices of that group and an overly critical view of other groups. Reversal is a negatively judging approach to evaluating the values and practices of one's own group and an uncritical view of those perceived as "other." Patricia displays characteristics of polarization in her constant comparison of her answer to Mark's answer and to Dr. Denton's verbal cues. Patricia focuses on the differences between the answers.

Transitioning to the next level of development involves noticing commonalities beneath the surface differences, in particular a growing awareness of norms. This middle orientation is "minimization," a lens for experience based on the idea, "Despite some differences, we really are all the same, deep down," and attends to similarity and universals (e.g., biological similarities – we all have to eat and sleep; and presumed universal values – we all know what good and evil are and the difference between them). The minimization orientation will, however, be blind to deeper recognition and appreciation of difference (e.g., Figure 2b, literally a "colorblind" view, what someone who has red/green colorblindness perceives). While feelings of sympathy (sorrow or joy for the experiences of someone else) are possible in polarization, a minimization orientation will tend to ethnocentric views that are not relationally dependent. For example, empathy might be confused for sympathy – a person with a largely minimization orientation may perceive feelings to be shared or common with a group of people without attempting an external validation of the perception with a relationally appropriate source (e.g., multiple members of the group). Mark's approach shows him seeking out ways in which his answer is similar to Patricia's and Dr. Denton's. A focus on commonality can bring everyone to a feeling of shared understanding but ignores subtle differences. In mathematics, this can lead to several potential problems. For instance, two people working together on a problem may have the same mathematical idea in mind but may not communicate effectively about the idea because of where each person focuses communication effort (e.g., on what is identical to their thought, or what is correct, or what is similar). As another example, at a meta-cognitive level, inattention to nuance may mean the difference between being stymied because a problem situation has no commonality with previous experience and the risk-taking of conjecturing a new and successful solution strategy by putting together old approaches in new ways.

Through increased attention to nuance in the differences that exist within noticed commonalities, one begins the transition from a minimization orientation to the "acceptance" orientation. Here, the word "acceptance" is used in its socio-cultural sense – the action or process of consenting to receive (rather than its psychological one – believe or come to recognize as valid or correct). Someone with an acceptance orientation has both some mindfulness of self as having a culture and awareness of moving among multiple cultures (plural). While an acceptance orientation supports empathy, awareness of difference, and the importance of relative context, how to respond and what to respond in-the-moment of interaction with others is still elusive.

The transition to "adaptation" involves developing culture-general frameworks for perception and behavior shifts that are responsive to a full spectrum of detail in an intercultural interaction (e.g., the detailed and contextualized view in Figure 2c along with a concomitant awareness that one's own perceptions (inside the frame) are limited and the whole picture is bigger than what we perceive). Adaptation is an orientation wherein one may shift cultural perspective, without losing or violating one's authentic self, and adjust communication and behavior in culturally appropriate ways.

Figure 3 shows the five orientations of this intercultural competence framework along with the shifts that occur in the transition from one orientation to the next. The movement along the continuum is not direct or linear. Folding back to previous orientations (particularly in times of stress) is common. Also, the time spent in learning about self and others during transitions and folding back hold value in developing more lenses through which one can view culture.

Knowing one's orientation, or the normative orientation of a group, can inform K-12 teacher and collegiate teacher work. In particular, we are researchers in a university-based project made up of several programs for in-service secondary mathematics teachers. Participants in the project include in-service teachers in a "mathematics for teaching" masters program, expert in-service teachers in a teacher leadership program, collegiate instructors for these programs, and mathematics education graduate students and faculty who are researchers on the project. Here we report on our early work to identify and build intercultural competence.

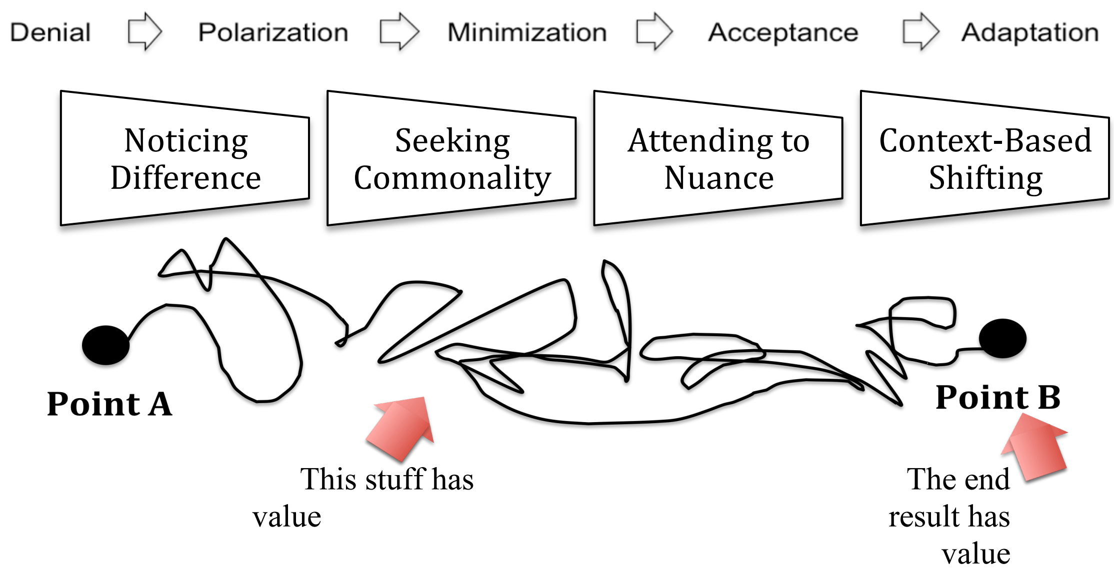


Figure 3. The transitions among stages in the continuum.

Research Question

What is the nature of intercultural competence development for mathematics instruction among university faculty and in-service secondary mathematics teachers participating in a university-based mathematics teacher professional development program?

Research Methods

Participants (26 in-service K-12 teachers and teacher leaders; 18 university faculty and grad students) completed the *Intercultural Development Inventory* (IDI), a reliable and valid method that identifies a person's intercultural orientation and elicits recent experiences and immediate cultural competence development goals (50 Likert-like items and 4 open response; Hammer, 2009). Each report from the IDI includes responses to the open-ended items along with quantitative information about *developmental orientation* (the orientation most likely at work in day-to-day interactions with others), *perceived orientation* (the orientation that a respondent perceives themselves to be working within) as well as *trailing orientations* (one or more fallback orientations likely to come into play in situations high in conflict or stress) and *leading* orientation (often aligned with perceived orientation, this is at the leading edge of someone's intercultural competence and the target for development). Two of the 4 open-ended items ask for respondents to tell stories: one involving an intercultural exchange that seemed to go well and one that did not go well.

From the IDI profiles we have a quantitative overview and, from the answers to the open-ended questions, material to help us in generating stories of intercultural challenges in teaching mathematics. The stories are the foundation for case study work with teachers and teacher educators. Our goal is cases that call up developmental, perceived, and leading orientations and provide space for discussing them and the transitions of awareness among them.

Results

In Figure 4 are the distributions among orientations for three groups who completed the IDI. As a group, the teachers' orientation was normatively in polarization while the teacher leaders, as a group, were largely at the lower end of minimization and the university folk were largely in minimization. As part of the research process, we conducted group profile debriefing sessions with teachers, teacher leaders, and university staff. When debriefing, three common goals emerged among all three groups of participants:

(1) build *awareness of self as having a cultural lens for viewing the world*;
(2) find guidance in the transitions through minimization and into acceptance, particularly how to *be mindful of one's cultural filter(s) for interacting with the world* (e.g., in the classroom, with colleagues, with other education stakeholders);
(3) engage in *building a knowledge base about equity*, including knowledge about culturally normative values and distinguishing these from essentializing or stereotyping approaches.

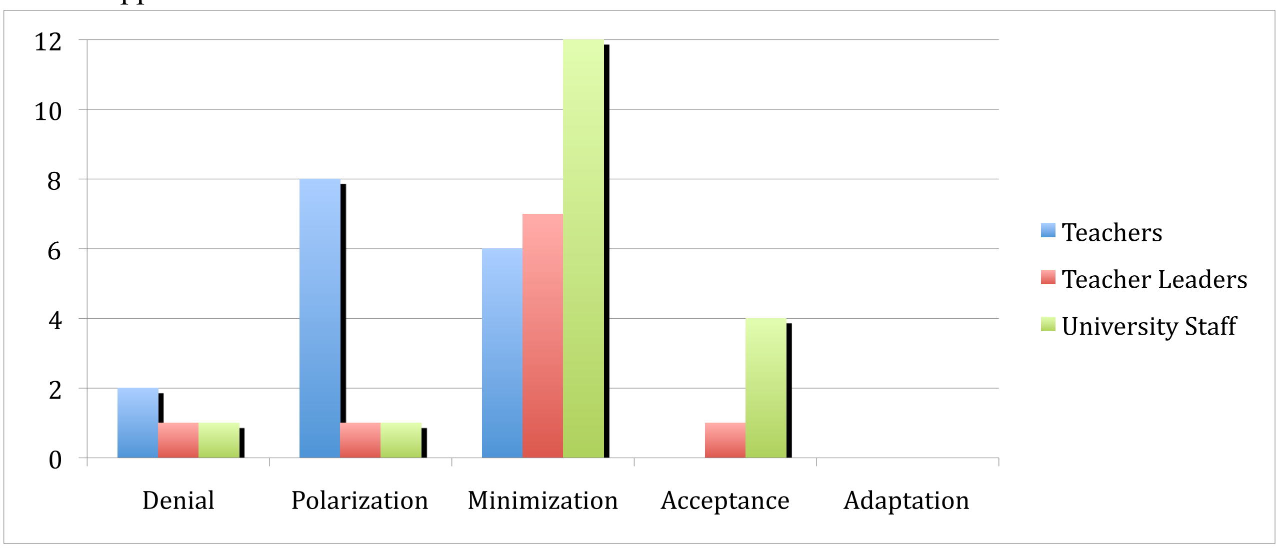


Figure 4. Distribution of Participant Developmental Orientations

Given the profile results, we see ourselves as having at least three different orientations for the case materials we are constructing: polarization, minimization, and acceptance. The statistical center of the teachers was at polarization. By putting the teacher character, Dr. Denton, in denial, we created a character who we were not expecting teachers or university staff to identify with, instead, they might recognize an earlier self (e.g., a trailing orientation that may arise during times of stress).

A case is not just a short story, it is a context-rich description in words, images, or both, of a dilemma, challenge, or epitome (e.g., authentic good or not-so-good practice). An effective case generates dissonance between what case users thought they knew to be true and what they are witnessing. Such cognitive dissonance is the basis on which new understanding is constructed. Associated with a vignette, to make a case, are prompts for the reader/viewer that depend on the content of the vignette, the degree to which it is experienced as intellectually/psychologically intricate, and the method of response (e.g., writing, discussing) (Seguin & Ambrosio; 2002).

Case prompts are especially effective if they draw the attention of the case user to four key areas of consideration and reflection:

1. *Framing*. Analyze different interpretations of the conflict, problem, or situation.
2. *Strategizing*. Evaluate the actions of the case participants and of oneself; consider how intentions are turned into actions in a variety of ways.
3. *Connecting*. Identify and relate personal experiences to the case experience.
4. *Forecasting*. Predict the consequences of actions, or inactions, for case participants and self-in-the-situation for the immediate and further future.

Prompts may call upon the reader to engage in complex synthesis, evaluation, and analysis of multiple sources of information, but can also be as simple as: Describe the problem, as you see it, in as much detail as possible. What might you do to deal with such a situation? Illustrate your strategy with specific examples from the vignette or personal experience. What, if any, would be the risks and the consequences of your strategy?

Building on What Occurred in the RUME 2011 Conference Session

Our goal at the conference was to share at least one potential-case situation with the audience and get feedback on possible IDI-based case prompts. That is, ideas for framing, strategizing, connecting, and forecasting in the context of the intercultural development continuum. In particular, we talked about the ways one might use the matrix in Figure 5 – an unpacking of the story about Dr. Denton, Patricia, and Mark – to generate and guide discussion that would grapple with difference, commonality, nuance, and context (i.e., the types of attention involved in the transition among stages, see Figure 3).

The example offered earlier in Figure 1 gives a description of the actions and comments made by Dr. Denton, Patricia, and Mark. Figure 5 below offers more detail for that scenario. The column on the left gives a description of the actions being performed by the two students. The three additional columns offer the scenario from each perspective. Bold font indicates utterances made, and the italicized font designates thoughts. This unpacking of the scenario allows for an in-depth look at the interactions of the intercultural competence orientations. As Patricia talks about how Mark's graph is different, Mark is noticing how it is similar but can be adapted to be the same as Patricia's graph. The bottom row offers alternative endings to the scenario showing the perspectives if each person operated in their leading orientation.

Session participants found the conceptual framework and consideration of Figure 5 useful in at least two ways. In the context of the case, it helped to organize their thoughts and perceptions of the case materials. Moreover, it helped them be reflective about themselves in situations in which they had participated. Additionally, several participants noted a feeling of recognition, that they saw themselves and saw colleagues in the descriptions (both those in the conceptual framework, Figures 2 and 3, and the illustrations in the characters in Figure 5). Several remarked that it was helpful in thinking about interactions and relationship building to pay attention to the transition activities: noting difference, seeing commonality, seeking nuance, and paying attention to the multiple orientations that may be participating in any interaction.

Description of actions while working on the prompt.	**Dr. Denton -** Discrete Mathematics Professor; Denial	**Patricia -** Pre-service Secondary Math; Polarization	Elem
Mark starts drawing vertices and connecting them to understand the problem.	*Why are you drawing?*	*But there's an equation, why would you do guess and check? [interprets drawing as guessing]*	**Let' sense. C it?**
Patricia writes Euler's Formula, V – E + F = 2, fills in values and solves.	*Yes, now you are doing it correctly.*	**Let's just use the equation like we are supposed to.**	*Oka and how picture.*
Mark and Patricia look at the prompt again. Mark looks at his drawing of a pentagon.	*They found there are 7 edges. There's really only one way to draw this, I should be able to go on to the next group soon.*	**Okay, I guess we do need to draw a graph.**	*Let's be a pen So I'll jo*
Patricia makes a drawing with 5 vertices and connects them with 7 edges.	**If you put the four dots in a square with a dot on top like a house with a roof, it will be easy to grade.**	*I could do it that way, but I put the last dot on the right of the square, because it made sense to me.*	*I tho Denton sure. I b*
Mark draws a new picture that looks like a "house," but two of the edges intersect.	*What you are doing does not make any sense.*	*But that's not right, we need 4 faces. That's got 5. Look, I already drew it right.*	**Do y**
Patricia and Mark look at each other's drawings while Dr. Denton is answering Mark's question.	**No. You have to have a roof.**	*Okay, what did you do wrong? Edges can't intersect!*	*How does she can see where is*
Mark erases line AD and instead connects A and B.	*Yes, that would make it correct.*	**Yours has 5 faces. Edges aren't supposed to intersect. You have to erase one of those in the middle and make a roof line.**	*Oh, differen like a ro*
Conclusion	**Yes, Mark's drawing is right.**	*So to make mine right, I would need to rotate it.*	*So r and hers Mine's j really th graph w edges.*
Alternate ending	*Technically you are correct, but I would prefer it drawn my way. [This represents a leading orientation of polarization.]*	*I like the one I drew; it's just rotated. It's not that different. [This represents a leading orientation of minimization.]*	*I wo there ar how you [Thi orientati*

Figure 5. Matrix of actions, thoughts, utterances that unpacks the interaction in Figure 1.

Implications and Applications for Research and Practice

What help in transitioning to global and ethnorelative mindsets can a teacher educator offer teachers (and teachers offer to students) if their own developmental orientation is more monocultural than intercultural? The challenge for any instructor is: how do I teach so that all students have opportunities to learn (not just the students with whom I experience cultural or orientational alignment)? The question applies to researchers as well: how do I research so that I get the perspectives of others (who may have a different orientation from my own)?

One way of addressing these questions, as theory developers, is through the nascent efforts reported here aimed at professional characteristic 5: we start with cases grounded in teacher reports of their classroom realities. On the other side of the same coin, for practitioners, we attend to theory by anchoring case activities (e.g., the points for discussion) in intercultural competence development for mathematics instruction.

Most mathematics educators juggle two identities, as problem-solver – one who can do mathematics – and as one who can teach it. So, in terms of the IDI, two future research possibilities occur to us. One is asking teachers and professors to take the IDI with a focus on *mathematics* as the culture being assumed as primary in answering each item. Then, having completed the survey wearing the "math hat," in order to cultivate information for more mathematical scenarios such as the one in Figures 1 and 5, it might be beneficial to have a question at the end of the IDI that is explicit in asking for connections between culture and mathematics in the classroom. For example, a story such as the example used here could be provided and a prompt would ask respondents to (1) comment on the story and (2) offer their own story. This would help to build further professional development scenarios and materials.

We are now revisiting our protocols. After all, if researchers have a minimization orientation where similarities are central, do research questions and instruments adequately capture the views and practices of teachers whose core communication about practice is focused on differences from a polarization orientation (or vice versa)? Similarly, when conducting research using observation protocols, the intercultural orientation(s) of protocols shape interpretations and frustrations with data gathered. We continue to learn about difference, commonality, and nuance ourselves as researchers.

Questions for the Reader

1. Consider the story of Helen in the appendix. What kinds of refocusing might we do to foreground the intersection of the culture of secondary mathematics, of textbooks, and of a teacher, with the framing of an assignment?

2. In an editorial, Ball, Goffney, and Bass (2005) have argued that in addition to teachers being culturally aware, that it is important for students to build adaptive competence in the culture of mathematics:

> In a democratic society, how disagreements are reconciled is crucial. But mathematics offers one set of experiences and norms for doing so, and other academic studies and experiences provide others. In literature, differences of interpretation need not be reconciled, in mathematics common consensus matters. In this way, mathematics contributes to young people's capacity for participation in a diverse society in which conflicts are not only an inescapable part of life, but their resolution, in disciplined ways, is a major source of growing new knowledge and practice. … Important to our argument is that these skills and practices that are central to mathematical work are ones that can contribute to the cultivation of

skills, habits, and dispositions for participation in a diverse democracy.

How might this perspective need to be revised or framed to be accessible to a teacher with a denial orientation? A polarization orientation? A minimization orientation?

3. In what ways might the cases we have discussed be useful OR need to be revised to be productive with pre-service teachers?

Acknowledgements

Thank you to the session participants at the RUME 2011 conference and to the organizers for their work in support of the conference and proceedings. This material is based upon work supported by the National Science Foundation under Grant No. DUE0832026 and U.S. Department of Education FIPSE Grant No. P116B060180. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the U.S. Department of Education.

References

Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups* (Tech. Report; NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Bennett, M. J. (2004). Becoming interculturally competent. In J. Wurzel (Ed.), *Towards multiculturalism: A reader in multicultural education* (2nd ed., pp. 62–77). Newton, MA: Intercultural Resource Corporation.

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (2nd ed., pp. 21-71) Yarmouth ME: Intercultural Press.

Bennett, J. M. & Bennett, M. J. (2004). Developing intercultural sensitivity: An integrative approach to global and domestic diversity. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of intercultural training* (pp. 147–165). Thousand Oaks, CA: Sage.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8).

DeJaeghere, J. G., & Cao, Y. (2009). Developing U.S. teachers' intercultural competence: Does professional development matter? *International Journal of Intercultural Relations*, *33*, 437-337.

Gay, G. (2000). *Culturally responsive teaching: Theory, research, and practice.* New York: Teachers College.

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education 53,* 106-116.

Greer, B., Nelson-Barber, S., Powell, A., & Mukhopadhyay, S. (2009)(Eds.), *Culturally responsive mathematics education*. Mahwah, NJ: Erlbaum.

Hammer, M. R. (2005), The Intercultural Conflict Style Inventory: A conceptual framework and measure of intercultural conflict approaches, International Journal of Intercultural Research, 29, 675–695.

Hammer, M. R. (2009). Solving problems and resolving conflict using the intercultural conflict style model and inventory. In M. A. Moodian (Ed.), *Contemporary leadership and intercultural competence: Exploring the cross-cultural dynamics within organizations* (pp. 219-232). Thousand Oaks, CA: Sage.

Keigher, A. (2010). *Teacher Attrition and Mobility: Results From the 2008–09 Teacher Follow-up Survey* (NCES 2010-353). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved February 28, 2011 from http://nces.ed.gov/pubsearch.

Kramsch, C. (1998). *Language and Culture*. New York: Oxford University Press.

Leininger, M. (2002). Culture care theory: A major contribution to advance transcultural nursing knowledge and practices. *Journal of Transcultural Nursing 13*(3), 189-192.

Seguin, C. A., & Ambrosio, A. L. (2002). Multicultural vignettes for teacher preparation. *Multicultural Perspectives, 4*(4), 10-16.

Shulman, L. S. (1998). Theory, practice, and the education of professionals. *The Elementary School Journal, 98*(5), 511-526.

Strizek, G.A., Pittsonberger, J.L., Riordan, K.E., Lyter, D.M., & Orlofsky, G.F. (2006). *Characteristics of Schools, Districts, Teachers, Principals, and School Libraries in the United States: 2003-04 Schools and Staffing Survey* (Tech. Report; NCES 2006-313 Revised). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

U.S. Army Training and Doctrine Command Culture Center, *Culture Education and Training Strategy for the U.S. Army*. Fort Huachuca, AZ: U.S. Army Intelligence Center, 2007.

Wolfel, R. (2008). *Culture Cubed: Towards Three-Part Definition of Intercultural Competence* (Tech. Report). Center for Languages, Cultures, and Regional Studies Position Paper, U.S. Military Academy. West Point, NY. Retrieved February 20, 2009 from <www.dean.usma.edu/centers/clcrs/ >

Appendix: Helen's Story

Example Case Story. Helen is a public high school mathematics teacher in a socio-economically and culturally diverse community. She is teaching a consumer mathematics class with mostly seniors. Helen wants all her students to believe they have what it takes to succeed in college so she has each student create a personal career portfolio. The assignment asks students to choose a job and a place to live after college. The portfolio is a report of research about living and working in this potential future career: starting pay for the job in that location, education required for that job, the cost of living in that location – which includes creating a budget for housing, utilities, transportation, food, and leisure. Included in the grading rubric are points for turning in a rough draft. Helen's intention is to provide three opportunities for students: (1) to see themselves as college graduates (2) to work with real-world numbers in creating a budget, (3) to receive feedback on a draft, with the expectation that the final report will have a higher score. Helen asks the class how the assignment is going and several student express frustration and confusion. She announces, again, that she will be available after school to help and is disappointed that students do not take advantage of this opportunity. Helen gets frustrated when several students who are not doing well already do not turn in a draft and do not come for help. She thinks to herself "If the students are struggling, why aren't they coming to my room for help?!" In speaking to one of her colleagues, she mentions her frustration.

[Pause here and discuss what elements of the transition from polarization to minimization might help Helen, what questions might need to be asked (and why) along with what advice Helen might be ready to hear and act on for refocusing of her attention in the situation.]

Helen's colleague Lee offered her own experience from high school, explaining that

> "going to office hours" in her middle school was as a form of punishment for misbehavior or low grades. In her first year of high school, the idea of going to office hours voluntarily made no sense to her: "Why would someone purposely take what amounted to an oral exam? Just to let the teacher know what she did not know and then be criticized for not knowing it?" Helen's first reaction was to dismiss Lee's story. "That's not what my office hours are like, that's not what *I* do!" Lee nodded and said, "Yes, I know. But I'm not completely sure how I learned that what it meant in high school to seek help from a teacher could be different from what it meant in middle school. In fact, the first time I went to an office hour in *college* it was because I was invited with two other people to have coffee in Professor Bladen's office – it was his sneaky way of getting us to the office so we could see what an office hour was like. And I've heard students talk about different reasons for not going to get help from teachers – like having a job during or working with parents or friends instead or because there was difficulty communicating with the teacher. So, I'm not sure why your current batch of students is not coming to your office, but there are probably lots of good reasons. Good to them, I mean." Helen shook her head, "That's too bad. Students should feel comfortable going to the teacher for help. Well, I can't help them if they don't come to see me. And, they won't come see me."

In the given story, Helen has a developmental orientation of polarization–defense. When working within a polarization orientation, what constitutes an "opportunity" is often decided with little or no consultation with the potential beneficiaries about whether it is seen as an opportunity. It could be that some of what Lee suggests is true, or that students in Helen's class were uncomfortable with her seeing their development process, or something else entirely. Discuss, again, what elements of the transition from polarization to minimization might help Helen, what questions might need to be asked (and why) along with what advice Helen might be ready to hear and act on for refocusing of her attention in the situation.

WHAT DO WE SEE? REAL TIME ASSESSMENT OF MIDDLE AND SECONDARY TEACHERS' PEDAGOGICAL CONTENT KNOWLEDGE

Billy Jackson, Saint Xavier University, bjackson@sxu.edu
Lisa Rice, University of Wyoming, lrice4@uwyo.edu
Kristin Noblet, U. of Northern Colorado, kristin.noblet@unco.edu

The article reviews efforts to develop an observation protocol to assess the pedagogical content knowledge (PCK) that middle and high school teachers may develop and demonstrate in the classroom over time as part of their participation in a master's program for secondary mathematics teachers. We observed each of 16 teachers in real time using the instrument, before involvement in the project and again during the second semester of participation. Aspects of the protocol measure four critical components of PCK including curricular content, discourse, anticipatory, and implementation knowledge. We present a quantitative analysis of the observations and discuss various challenges faced in the instrument development and its relation to similar protocols used by others previously.

Key Words: pedagogical content knowledge, discourse, inter-rater reliability, teaching moves

There have been several approaches to measuring the pedagogical content knowledge (PCK) of practicing teachers. Indeed, Hill, Ball, and Schilling (2008) and Hauk, Jackson, and Noblet (2010) have documented their development of written instruments designed to assess aspects of PCK. Both groups have developed theoretical frameworks for PCK that have similarities and some differences. One of the principle differences is that the Hill, Ball and Schilling model seeks to measure each of their proposed categories of PCK as distinct from each other, while Hauk, Jackson, and Noblet take a non-linear approach that presumes measurement overlap among categories. To understand what a non-linear model means, PCK can be thought of as a dynamical system as its development in teachers progresses over time. Hill, Ball, and Schilling presume that for their written instrument, a teacher's PCK can be measured by taking a linear combination of items that each pertain to exactly one of their three PCK constructs: knowledge of content and students (KCS), knowledge of content and teaching (KCT), and knowledge of content and curriculum (KCC). In contrast, a single item on the written instrument from Hauk, Jackson, and Noblet can touch on multiple facets of their typology of PCK, and hence introduces overlap in these facets, creating nonlinearity in the system. While Ball and colleagues acknowledge that the relationship between a teacher's KCS, KCC, and KCT may well be non-linear (LMT, 2006), the current model makes the linearity assumption. A mathematical model of PCK with consideration of Hauk, Jackson, and Noblet's approach will require a non-linear system.

Hauk, Jackson, and Noblet (2010) discuss PCK in terms of four components: curricular content, discourse, anticipatory, and implementation (action) knowledge. *Curricular content knowledge* is "substantive knowledge about topics, procedures, and concepts along with a comprehension of the relationships among them *as they are offered in school curricula*" (p. 2). *Discourse knowledge* is substantive knowledge "about the culturally embedded nature of inquiry and forms of communication in mathematics (both in and out of educational settings" (p.2), and as such includes knowledge of mathematical syntax as a sub-category. *Anticipatory knowledge* "is an awareness of, and responsiveness to, the diverse ways in which learners may engage with content, processes, and concepts" (p. 3). *Implementation* or *action* knowledge "includes knowledge about how to adapt teaching according to content and socio-cultural context and *enact in the classroom* the decisions informed by content, discourse, and anticipatory understandings" (p. 3).

Both groups' written assessments use multiple-choice items and are limited in measurement of teacher knowledge related to implementation. Such knowledge is more challenging to assess as this type of knowledge requires actions executed in the classroom (e.g., teacher moves). That is, the written assessments could not test for this type of knowledge because it requires that the teacher act "in the moment." At best, any written item could only gauge what a teacher *might* do in certain situations (e.g., see Ball, Hill, & Schilling, 2008). In order to validate their written instrument, Ball and others (Learning Mathematics for Teaching (LMT), 2006) developed another instrument aimed at quantitatively measuring aspects of elementary and middle school teachers' classroom practice. Ten K-8 teachers who had taken a PCK test were videotaped three times prior to, during, and after participation in professional development. Over the course of a year, a team of mathematicians, mathematics educators, mathematics teachers, and non-specialists analyzed the videos for various aspects of mathematics and mathematics teaching present in each lesson. A rubric was developed containing several items and video reviewers trained for and then coded each 5 minute segment of each lesson for different categories of teacher move or classroom interaction. Each category had four possible codes: Present and Appropriate (PA), Present and Inappropriate (PI), Not Present and Appropriate (NPA) and Not Present and Inappropriate (NPI). LMT team leaders noticed early on a wide variability in how individuals coded lessons based upon the individuals' own professional backgrounds, and so to help ensure inter-rater reliability, the lessons were all recoded in pairs. A glossary describing each category (column) in the observation rubric was written, with each description giving some detail on when each code should be assigned during a segment.

Theoretical Perspective

Our research blends the Hauk et al., framework for PCK and the LMT observation instrument designed by the research team at the University of Michigan. We take the view that the teacher actions or moves (or the absence thereof) in the LMT protocol can be observed in the classroom, and that such actions or moves can be described (at least approximately) in a predetermined coding format *independently of the researcher involved*. Now, this is not to say that two different researchers may not observe and record different things (as frequently happened with the team at the University of Michigan and for our team) for a given segment a few minutes long, but, like the LMT tool, for an entire class period observation we would expect variation between observer totals to be minimal.

We use here the typologies of Hauk, Jackson, and Noblet (2010). The reason is that any particular move that a teacher makes in the classroom can demonstrate multiple facets of PCK simultaneously, and hence we take their view that the strands of PCK are interwoven during instruction. Also, Hauk, Jackson, and Noblet make cultures in the classroom an explicit part of their definitions, which in turn may be part of teacher decisions to make certain moves in response to them.

Discourse knowledge requires some further investigation here, as it is necessary to make explicit how we are defining it. Ryve (2011) points out that it is indeed important for authors to provide definitions and epistemologies when discussing discourse as "some key words or phrases are particularly complex and multifaceted" (p. 167). To this end, we use Gee's (1990) definition of Discourse (with a capital 'D'), which he distinguishes from discourse (with a lower case 'd').

For Gee, “Discourses are ways of being in the world, or forms of life which integrate words, acts, values, beliefs, attitudes, social identities, as well as gestures, glances, body positions and clothes” (pp. 142-143). However, “discourse” with a little “d” refers to “connected stretches of language that makes sense” (pp. 142-143). Hence for Gee, discourse is a part of Discourse. Note that if discourse is "connected stretches of language that make sense" to those taking in and/or producing the connected stretches, then semantics and syntax are part of discourse. That is, discourse could be thought of to have at least three components: what is intended, what is perceived, and the associated interaction that allows the exchange from intender to perceiver. Such discourse applies to written/read and or spoken/heard interaction. For the purposes of this paper, whenever the word “discourse” is used, we will mean the word in the broader sense of Gee’s “Discourse” as such things as gestures will be considered discourse for our purposes, though, admittedly, here usually we will refer to spoken language. The discourse used here is that of mathematical discourse rather than that of what Ryve (2011) terms “general educational discourses” (p. 174) that can be relevant to mathematical discourse but are nonmathematical in nature.

Ryve also points out that our epistemological assumptions have consequences for how we conceptualize mathematics and learning. In light of this, we take Sfard’s (2006, 2008) view that mathematical objects are concrete and tangible and hence can be perceived by the learner. Examination of the LMT (2006) instrument and the corresponding glossary shows that Ball and others share this same view in their work as well.

Discourse knowledge is perhaps the most unique type of knowledge among the classifications of Hauk, Jackson, and Noblet. A comparison of discourse knowledge to curricular content knowledge reveals that curricular content can become quite static at some point while discourse knowledge constantly evolves due to the nature of discourse being imbedded in the culture of the classroom, which is subject to change on both short and long time scales. In fact, many shifts in PCK that occur as a teacher develops it can be attributed to changes in discourse knowledge, since discourse is woven into the other three components very deeply (i.e. there is substantial overlap between discourse and the other three). Discourse knowledge, like implementation knowledge, might have aspects of social co-construction whereas curricular content and anticipatory knowledge may tend to be individually constructed (though such construction can be socially mediated).

The research questions for the work reported here are: How might we track the effects of professional development through changes in observed PCK? If traceable, how might professional development be designed to foster particular classroom moves through changes in PCK? Work on both of these questions continues, and we will primarily address the first here but some attention will be given to implications of current results for the second.

Methods

We built our work from a careful consideration of the LMT work. The LMT group point out in their technical report that there is a need to develop an instrument for doing observations in real time (LMT, 2006, p. 20). In order to address this need, we examined their observation protocol in some depth and determined which items were most appropriate given our focus on observing in secondary mathematics classrooms in real time (the LMT work was in grades K-8). A copy of our protocol is in the appendix. Their protocol contained over 30 categories. To streamline for real-time observation we shortened to a protocol containing 20 items. Some of their categories were replaced or condensed in our version. For example, in the LMT version, the researchers created columns for the following: selection of correct manipulatives, and other visual and concrete models to represent mathematical ideas (their column II.e on sheet 2) and multiple models (column II.f on sheet 2). In our version, these two columns were condensed into the column that we titled multiple representations, which could include all of the things that the LMT team was looking for in II.e and II.f.

Great care was taken in finding an appropriate length of a segment to be viewed during the class. The team started with the 5 minute length that the LMT used for recorded sessions, but it soon became clear that a "5 minute on, 5 minute off" strategy in which the researcher would observe for 5 minutes and then record tallies on the protocol during the next 5 minute interval would result in 5 or fewer codings per class period for each category (not to mention the fact that it was challenging to hold five minutes of observation in one's head while attempting to record everything about it). Eventually, the team agreed upon observing for 3 minutes, and then recording for 3 minutes. Thus, it took 6 minutes to capture a particular 3 minute chunk of observation on the observation sheet, which meant about 9 samples in a short class period (45 to 55 minutes) and up to 15 in an extended or block period.

After the team started using the protocol, we began to reexamine the glossary that the LMT team had developed. We found that trying to use the instrument in real time created new challenges with respect to inter-rater reliability. In particular, the words "explicit" and "inappropriate" leave much room for interpretation even in the definitions provided by the LMT team. Though we used many of the same column categories and indentifying language as they did, we also saw it was important to craft definitions and create a new glossary. The idea was to create an instrument with sufficient examples and non-examples for each category that it could act as a coding book: a guide to the intended viewpoint of the protocol and how to observe through a particular lens. The eventual goal is to have an instrument that is terse but of sufficient detail that individuals can observe classrooms after a brief calibration training paired with a practiced observer or after watching video of the same.

For example, while our glossary continues to be refined, we felt a need to be, well, more explicit about what "explicit talk about a topic or subject" meant. Currently, our glossary description of this category is: any utterance from student or teacher in which a topic or subject is stated verbally or in writing or by reference to a clear written or verbal precursor familiar to people in the room (e.g. through gestures). In vivo exemplars have been included in our glossary to demonstrate categories. For example, during one 3-minute segment, the teacher presented the Fundamental Theorem of Algebra. The exercise the teacher assigned called for students to find a polynomial of lowest degree with real coefficients that had certain prescribed roots. At one point, an exercise asked for a polynomial with roots $3i$, 4, and 5. The teacher produced a monic degree 4 polynomial with these 3 prescribed roots, and a student asked why it was necessary to have $-3i$ as a root when this number was not contained in the list. The teacher responded that since $3i$ was a root, its conjugate $-3i$ also had to be a root. The student again asked why this must be true when $-3i$ was not listed, and the teacher replied "because conjugates are always roots." The researcher coded this particular segment as NPI in the "explicit talk about ways of reasoning" column due to the teacher's not addressing directly the student's question in a way that appeared to make sense to the student. The segment was also coded as NPI in the "interprets students' productions/student errors" column due to the fact that in the glossary, this column includes the teacher listening to what the student says and responding to the student.

Each column of the protocol was assigned a quadruplet of the form (c, d, a, i) where the values of c, d, a, i were determined by the research team based on the descriptions of the categories **C**urricular Content, **D**iscourse, **A**nticipatory, **I**mplementation and the glossary description of the category represented by the column. The value for *c, d, a,* or *i* was 1 if a particular kind of knowledge was present in the observable category, or 0 otherwise. That is, we would say a particular category "loaded" on a particular PCK aspect if the appropriate presence of that category suggested that the teacher demonstrated that component of PCK during the segment observed. Research team members spent a significant amount of time on coming to a consensus on each of the loadings, particularly implementation knowledge. Working to understand how in what ways the observation categories were aligned to the idea of action knowledge being demonstrated in the classroom. One challenge in defining this particular type of knowledge is that the other three are interwoven with it so much that at times it can be difficult to "tease apart" implementation knowledge from say anticipatory knowledge. After much discussion, we began to understand that implementation knowledge had to meet **both** criteria given in the definition by Hauk, Jackson, and Noblet (i.e., adapting and enacting). This categorical inductive coding left some of the columns without non-zero alignment to any PCK codes (e.g. instructional time spent on mathematics(>75%)). One particularly interesting column titled "encourages diverse mathematical competencies" has a unique feature: we determined that this column loaded heavily on PCK by assigning it a quadruple of (1,0,1,1). The category draws on curriculum content knowledge because the teacher will demonstrate some knowledge of the curriculum even if peripherally since the teacher must be informed about the curriculum enough to know that she is encouraging diverse mathematical competencies. She is also demonstrating the use of her anticipatory knowledge when the column is checked as PA because she must be aware of how the students may interact with the material at hand in order to encourage the competencies. She adapts her teaching according to the diversities that arise and makes choices in her instruction accordingly, and so the observer will see her use her discourse knowledge during the segment.

The columns that have discourse knowledge present share a feature the definitions of both take into consideration the *interaction* that occurs between the teacher and students as opposed to solely the *actions* of the teacher. Hence, another difference between the typologies of Ball, Hill, and Schilling and Hauk, Jackson, and Noblet becomes evident. The approach of the latter attends to the relational and interactional. Thus, our instrument and the lens it takes may allow for the researcher to look for certain aspects of the classroom that the LMT instrument does not.

Results

The research team observed two cohorts of teacher participants (TP) for a masters program for teachers (n=19 and n=16 respectively). We observed both cohorts before entering the program and the first cohort during their second semester in the program. A frequency analysis was performed for each observation and t-tests performed.

The frequency analysis began by summing the PA's and PI/NPI's for each column. We then subtracted the number of inappropriates from the number of appropriates, assigning a value of +1 for each segment checked PA, -1 for each segment checked PI/NPI, and a value of 0 for each segment marked NPA. After performing this tally for each column, each of the column totals in the protocol were normed in each observation to account for varying numbers of segments coded among the 2 to 3 observations done for each teacher. A PCK score for each of curricular content knowledge (CCK), discourse knowledge (DK), anticipatory knowledge (AK) and implementation knowledge (IK) was assigned by summing the columns for which the team assigned a value of 1 in the quadruplet. For example, the column "encourages diverse mathematical competencies'' contributes to the CCK, AK, and IK scores but not to the DK score since the column was assigned a quadruple of (1,0,1,1) as explained in the previous section. The normed PCK scores from each observation were then averaged and the teacher was assigned that average in each component for pre and follow up observations.

A paired *t*-test was then done for the first cohort on the pre and follow up observations, and an independent samples t-test was done for comparison between the first and seconds cohorts' pre-observations. The results are summarized in the tables below:

Type of Knowledge	*Mean*	*p-value*
CCK, pre, cohort 1	.5926	
CCK, follow up, cohort 1	.5867	.942
DK, pre, cohort 1	.4360	
DK, follow up, cohort 1	.4787	.394
AK, pre, cohort 1	.6239	
AK, follow up, cohort 1	.5098	.03
IK, pre, cohort 1	.5188	
IK, follow up, cohort 1	.5051	.774

Type of Knowledge	Mean	p-value
CCK, pre, cohort 1	.593	
CCK, pre, cohort 2	.605	.818
DK, pre, cohort 1	.436	
DK, pre, cohort 2	.6214	.000
AK, pre, cohort 1	.624	
AK, pre, cohort 2	.588	.459
IK, pre, cohort 1	.519	
IK, pre, cohort 2	.588	.081

Cohort 1 showed no significant differences in the pre and follow up observations in any of the categories except for anticipatory knowledge, in which a significant decrease was observed. This decrease could be due to any number of factors, including an implementation dip. Comparing cohorts 1 and 2, we see that cohort 2 had statistically significantly different DK and IK scores on arriving in the master's program. To understand what these means represent, we first note that inappropriates were assigned infrequently during the observations as researchers tended to describe the classes as average to above average. Thus, a mean of .519 for cohort 1 IK implies that the teachers in cohort 1 demonstrated the use of their IK in about 51.9% of the segments observed, while cohort 2 demonstrated the use of their IK in about 58.8% of the segments observed. It is important to note that this does not mean that cohort 2 has more IK than cohort 1 as a teacher can have knowledge but not use this knowledge during any particular instructional segment. The statistical differences along with other external factors could suggest that the two cohorts are different as populations.

It is interesting to note the difference in discourse knowledge between the to cohorts. It is possible that the discourse knowledge is a factor in the difference in the implementation knowledge between the two cohorts due to the overlap between these two components of PCK. One possible implication of this is that designers of professional development (PD) who wish to create shifts in PCK may be effective by targeting teachers' discourse knowledge. Hence, for example, to advance a teacher's anticipatory knowledge, a PD designer may create sessions designed to build upon her discourse knowledge since these categories overlap in PCK framework.

Conclusions

The model for PCK offered by Hauk, Jackson, and Noblet has some striking differences from that of Hill, Ball, and Schilling. The instrument our research team developed based upon the LMT project in many ways highlights many of those differences. Perhaps one of the biggest strengths of using the nonlinear model is that it foregrounds the importance of discourse in the mathematics classroom, activity that creates interaction among the teacher and students rather than being a sole action of the teacher. It is also possible that discourse knowledge may be responsible for shifts in PCK due to its intrinsic overlap with the other components that Hauk, Jackson, and Noblet propose. This realization could have consequences for the design and implementation of professional development aimed at targeting PCK. We plan further research in this area to examine how PD can be designed with explicit attention to discourse and ways to "unpack" what it means to talk about Discourse (and discourse) in the classroom, particularly in classrooms where students and teachers from myriad cultural Discourse communities interact.

Acknowledgements

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

Gee, J. P. (1990). *Social linguistics and literacies: Ideology in discourses* (1st ed). London: Falmer Press.

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

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.

Learning Mathematics for Teaching (2006). *A Coding rubric for Measuring the Mathematical Quality of Instruction* (Technical Report LMT1.06*)*. Ann Arbor, MI: University of Michigan, School of Education.

Ryve, A. (2011). Discourse research in mathematics education: A critical evaluation of 108 Journal Articles. *Journal for Research in Mathematics Education,* 42, 167-198.

Sfard, A. (2006). Participationist discourse on mathematics learning. In J. Maasz & W. Schloeglman (Eds.), *New mathematics education research and practice* (pp. 153-170). Rotterdam, The Netherlands: Sense.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge, UK: Cambridge University Press.

Appendix-Observation Protocol

Math TLC Teaching Observations: Recording Sheet

Teacher ______ **Course** ______ **Date** ______ **Time** ______ **OBS #** ______

Directions: In all but A., B., and C.a., choose only one option for each class segment.

I. Instructional Format and Practices

	A. Format for segment			B. Lesson/segment type					C. Mathematics Teaching Practices																															
	a. whole group	b. small group/ partner	c. individual	a. review, warm up or homework	b. Introducing major task	c. Student work time	d. Direct instruction	e. Synthesis or closure	a. Voices in the classroom (students, teachers, or both)				b. Real-world problems or examples				c. Interprets students' productions/ students' errors				d. Explicit about student tasks, work, and success				e. Explicit talk about mathematical practices & language				f. Explicit talk about ways of reasoning				g. Instructional time is spent on mathematics (>75% of segment)				h. Encourages diverse mathematical competencies			
									Students		Teachers		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
Segment									A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
1																																								
2																																								
3																																								
4																																								
5																																								
6																																								
7																																								
8																																								
9																																								
10																																								
11																																								
12																																								
13																																								
14																																								
15																																								
16																																								
17																																								
18																																								
19																																								
20																																								

Math TLC- Teaching Observation Recording Sheet

II. Knowledge of mathematical terrain of enacted lesson

DATE: ________	a. Conventional notation (mathematical symbols)				b. Technical language (mathematical terms and concepts)				c. General language for expressing mathematical ideas (overall care and precision with language)				d. Selection of numbers, cases & contexts for mathematical ideas				e. Uses multiple ways to represent mathematical ideas				f. Makes links/connections among mutliple representations of mathematical ideas				g. Publicly records the mathematics for the class				h. Mathematical descriptions (of steps)				i. Mathematical explanations (giving mathematical meaning to ideas or procedures)				j. Mathematical justifications				k. Development of mathematical elements of the work (i.e., moving the math along)				l. Computational errors or other mathematical oversights			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
Segment	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
1																																																
2																																																
3																																																
4																																																
5																																																
6																																																
7																																																
8																																																
9																																																
10																																																
11																																																
12																																																
13																																																
14																																																
15																																																
16																																																
17																																																
18																																																
19																																																
20																																																

Overall level of teacher's knowledge of mathematics

Low	Medium	High

Math TLC- Teaching Observation Recording Sheet

WHAT DO LECTURE TEACHERS BRING TO A STUDENT-CENTERED CLASSROOM? A CATALOGUE OF LECTURE TEACHER MOVES

Estrella Johnson, Carolyn McCaffrey
Portland State University

In this article we aim to understand what it looks like when community college instructors, with little to no experience with inquiry-oriented curriculum, implement inquiry-based curriculum for the first time. To approach this question we focused our attention on a community college instructor's first implementation of an inquiry-oriented task. In total we identified six moves used during the implementation of this task. These moves are (1) zooming out, (2) real world examples, (3) counter-examples, (4) selective restating (5) referring to definitions, and (6) sequencing of student sharing. By identifying these teacher moves, it is indicated that mathematics instructors, even ones who primarily engage in teacher-center teaching, have techniques that they can draw on as they enact inquiry-oriented curriculum materials. Identifying such techniques can serve as a starting point for understanding how to support college-level teachers in changing their teaching practices.

Key Words: Inquiry-oriented, Teaching, Teacher Moves, Community College

Introduction

At the turn of the millennium, mathematics education reformers placed a renewed importance on inquiry-based instruction. The National Council of Teachers of Mathematics (2000) recommended that "teachers [should] help students make, refine, and explore conjectures," and that students should become "flexible and resourceful problem solvers". However, the pedagogical skills necessary for implementing inquiry-based tasks may differ from those of lecture based pedagogy. Although the needs of reform based instruction have been explored at the K-12 grades (Ball, 1993; Bowler, 1998, 2006; Cohen, 1990; Wood, 2001) and to a lesser degree at an undergraduate level (Rasmussen & Marrongelle, 2006; Speer & Wagner, 2009), very little research has been done on community college instruction of inquiry-based curriculum. Community colleges are unique environment, both because of the high numbers of non-traditional students and because of the large variability in teacher training and background. Questions about community college instruction become more pressing as enrollment continues to increase (Phillippe & Sullivan, 2006).

We sought to understand what it looks like when community college instructors, with little to no experience with inquiry-oriented curriculum, implement inquiry-based curriculum for the first time. To approach this question we focused our attention on a community college instructor's, Bill's[5], first implementation of an inquiry-oriented task. While Bill had been teaching one and two hundred level mathematics courses, such as pre-calculus and linear algebra, for four years at the community college (and prior to this taught at a high-school for twelve years), this course was Bill's first student-centered teaching experience.

[5] All names in this report are pseudonyms.

Study Context

Our study takes place within the broader context of a project aimed to develop a community college "transition to proof" course, based on an inquiry-oriented abstract algebra curriculum – *Teaching Abstract Algebra for Undersatnding* (Larsen, 2009; Larsen, et al., 2009; Larsen et al., 2011). The task we focus on was the very first mathematical investigation of the course. In this task students were initially given six shapes (see figure 1 below) and were asked to arrange the figures from least to most symmetric. The students worked on this task individually and then in small groups prior to a whole class discussion. The groups shared how they ordered the figures and how they came to that decision. The students were then asked to determine a way to quantify the symmetry of each figure and, using their quantification criteria, the groups ranked the figures and presented both their criteria and their ranking to the whole class. Following these presentations the groups worked to develop both a definition of what a symmetry is and what makes two symmetries equivalent (see Larsen & Bartlo, 2009), the learning goals of this task. In total, this task covered three classes, each one hour and fifteen minutes in duration. The majority of class time was spent with students working in small groups or presenting their ideas in a whole class setting; Bill's role in these classes was to launch the tasks, monitor group work, and facilitate whole class discussion.

1. (Private Think Time) **Rank the figures in order from least to most symmetric.** Note: there is not a single right way to do this. Base your responses on instinct, intuition, or aesthetic sense.

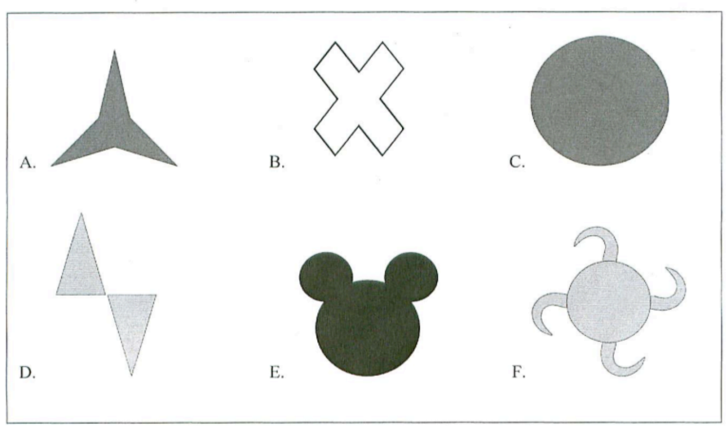


Fig. 1 Symmetry Task Launch

Research Question

Through initial analysis of the classroom video, we refined our research question to focus on the teacher moves Bill had at his disposal during this task. Here we use the term "teacher moves" to refer to specific actions that direct the mathematical trajectory of the lesson, such as providing counter-examples and sequencing student contributions. Specifically, we will examine:

> What types of teacher moves did Bill utilize to direct the trajectory of the lesson as he implemented an inquiry-oriented curriculum task for the first time?

In focusing narrowly on these moves, we are not suggesting that teacher moves exemplify the complexity of teaching practice, nor do we assert that gaining proficiency with specific moves is sufficient for effective teaching. However, understanding what moves teachers may naturally gain though teacher-centered, lecture classes can help inform why teachers struggle to implement reform curriculum and suggest starting-points for professional development and teacher support.

Related Literature

Given that Bill's primary classroom role during this task implementation involved leading and facilitating whole class discussions, we were curious about possible teacher moves that have been shown to advance the mathematical agenda during whole class discussion and student sharing. Other researchers have previously examined the ways teachers negotiate whole class discourse to promote student understanding. Rasmussen and Marrongelle (2006) have proposed a theoretical construct called *pedagogical content tools,* which they define as "a device, such as a graph, diagram, equation, or verbal statement, that a teacher intentionally uses to connect to student thinking while moving the mathematical agenda forward" (pg. 389). They describe and illustrate two such tools, *generative alternatives* and *transformational records* within the context of an undergraduate differential equations classroom.

Additionally, Speer and Wagner (2009) investigated some of the challenges faced by Gage, an experienced mathematics professor, in his first attempt to enact an inquiry-oriented curriculum in a differential equations course. The authors focused on *analytic scaffolding*, the teacher contributions that support the development of mathematical ideas for students. They found that deficiencies in Gage's pedagogical content knowledge led to an inability to create sufficient analytic scaffolding to support classroom discourse and to move the mathematics forward.

We approached our data set with an eye tuned toward teacher tools such as pedagogical content tools and analytic scaffolding, however we were also interested in identifying other teacher moves that both moved the classroom towards the intended goal of the task and contributed to student mathematical sense-making, such as questioning and sequencing. We use the term teacher moves, as opposed to tools, to reflect our focus on any specific teacher action that redirects the trajectory of lesson as it is unfolding in the moment. This contrasts with the formal notion of tools, which were defined by Rasmussen and Marrongelle (2006) as "something that the informed user explicitly recognizes as useful for achieving specific goals" (p. 389). Therefore, when looking for teacher moves we did not require that Bill explicitly recognized their usefulness or that these moves were done with a conscious intention.

Methods

All of the three class periods in which the symmetry task was implemented were video recorded, with the camera focusing on Bill during whole class discussion and on a single group of students during individual and group work. The research team made several passes through this video data, consistent with iterative video analysis (Lesh & Lehrer, 2000), in order to refine our initial research questions and then to identify teacher moves used by Bill to facilitate whole class discussions and facilitate student construction of understanding.

During the first pass of analysis, a video log was made, in which time stamps were recorded to identify the type of class activity (i.e. group work, poster presentations) and to make note of times in which Bill was leading/facilitating whole class discussion. Additionally, key mathematical moments of the task were identified, such as when the definition of symmetry was introduced to the class. It was during this phase of analysis that the research team began to recognize specific instances in which Bill stepped in to direct the trajectory of the lesson. For instance, when the students were initially given the figures and were asked to order them from least to most symmetric, some students were basing their decisions on how many times you could fold the figure such that the two halves would be the same. Upon hearing such reasoning during small group discussions, Bill redirected students to think of reflecting the figures as opposed to folding the figures[6]. In describing the mechanism in which Bill redirected the class away from folding, the research team became curious as to other moves used by Bill during the implementation of this task.

In the second phase of analysis, each member of the research team watched one or two of the three days of classroom video data, this time with the specific goal of identifying moves used by Bill to direct and facilitate whole class discussion. The research team then met to share possible moves and the video clips in which these possible moves were present. During this final stage of analysis the research team reached a consensus about both which moves were present and which clips were illustrative of these moves.

Results and Discussion

In total we identified six moves used by Bill over three days, during the implementation of this task. These moves are (1) zooming out, (2) real world examples, (3) counter-examples, (4) selective restating (5) referring to definitions, and (6) sequencing of student sharing. In this section we will describe each of these moves and provide examples of how Bill used these moves during whole class discussions.

Zooming Out

Twice during the implementation of this task, Bill made reference to "how things will be done in this class" or to "how the mathematical community works" as a way to motivate or justify a change in the trajectory in the lesson. One example of this was when Bill introduced the definition of a symmetry to the class. At this point in the task, each group had just presented their ranking of the figures and the criteria they used to quantify the symmetry of a figure. In an effort to motivate why Bill would give the class a definition of symmetry, Bill said the following to the class:

> Bill: You are going to be presented *in this course* with different problems or different puzzles. And often after a little playing around everybody is going to end up on pretty much the same track, because there is just not so many ways to approach them. Other times there really is a lot of different ways to approach it, and so there is no kind of structure that would kind of, put everyone in the same channel. When that happens, which will happen every now and then, *that's when I'm going to step in and present some formal definition, or something like*

[6] Bill was guiding the class towards a definition of symmetry as a type of rigid motion. Identifying symmetries in terms of "folds" is in conflict with the idea of rigid motion.

> *that*. *That will happen occasionally as we go through the course*. And if I were a student in your shoes, I would feel comforted with that, that I'm not inventing mathematics that is different from what everyone else is doing.

In this statement, Bill described a classroom norm: in times of classroom debate, Bill may step in to provide formal mathematical definitions. In this specific instance, Bill's use of referring to this norm served to justify the rapid change in trajectory of the class. Directly following this explanation Bill presented a definition of symmetry, stating that "a symmetry is a rigid motion, that when applied to a figure, results in the figure landing on itself". This definition of symmetry had been used by previous instructors of this curriculum. So, this definition was developed prior to Bill's implementation of this task and thus in absence of his students' emerging conceptions of symmetry. By introducing the definition of a symmetry to the class this way, as opposed to building off of the ranking criteria developed by the students, such as "the number of ways to reflect the figure to make it land on itself", Bill drew the attention of the class away from the students' work on the task and towards a more general description of how this course will operate. Thus, this serves as an example of how Bill *zoomed out* as a way to direct the trajectory of the lesson.

Real World Examples

Three times in the implementation of this task, Bill made use of real world examples to clearly illustrate mathematical ideas. In both of the following excerpts, the examples served as analogies for mathematical definitions. In the following transcript, Bill builds upon the student description of what it means for a figure to "land on itself." The student described "landing on itself" to mean that all of the boarders aligned on one another. Bill elaborated further with the following:

> Bill: I'm wondering, what's a good image that I could use that a lot of us are going to really connect with well. And the mentioning of boundaries made me think of, um, toys that toddlers and sub-toddlers play with. You know those toys that help toddlers figure out what's a circle and what's a square, and what's a star. And so it would be like some board, with a kind of die cut shapes, and try to fit the shapes in. That sounds like what you are describing.
>
> S1: Yeah.
>
> Bill: Does that give everyone kind of a common framework. I wanted to give some sort of vivid image to support what we have been doing.

In this passage it is clear that Bill wished to establish a common classroom conception of "landing on itself" by tying the idea to a real world child's game that he hoped was common to the students. This analogy was useful to at least one student. Later in the class period when students were discussing equivalence, one student made reference to "fitting the star into the star hole," which built upon Bill's initial imagery.

In the next example, Bill settled a class dispute about the distinction between the ideas of "equivalence" and being "the same." The following transcript begins with a student explaining the similarity between a 0 angle and a 2π angle.

S1: Is it [a zero angle] the same as a 2π angle?
Bill: Is it?
S2: How can you tell the difference?
S1: They have the same sine and the same cosine.
S3: But, he's asking if equivalent means the same.
S4: You aren't doing the same, but you end up with exactly the same.
S3: So it probably wouldn't be, because you are doing two different symmetries that are equivalent, but not the same.
Bill: *So if you performed a 360° rotation on yourself, versus a 720 degree rotation on yourself, it is equivalent, but is it identical?*
Students: No, no.
S1: The more you do the dizzier you get.
Bill: So is a 360° equivalent to a zero degree?
S5: No.
Bill: Equivalent?
S5: Oh, equivalent yes.
[other students say yes]
Bill: Identical?
Students: No.

In this passage, Bill used the real world example of a student physically spinning in circles to draw a clear distinction between "the same" and "equivalent." Through the use of the example, students were led to the mathematically conventional conception that equivalence is not the same as being identical, which is a crucial concept in abstract algebra.

It is important to note that this type of teacher move, introducing illustrative examples from a real world context, may be a valuable technique within a lecture-based classroom. In teacher-directed classes, teachers are wholly responsible for presenting material in a clear manner that is easily assimilated by the students. Real world examples can help students visualize or understand abstract material presented by the instructor. During the inquiry-based task, Bill carried over this technique to "settle" student disputes and to restate a student idea in a more vivid way.

Counter-Examples

Another technique that Bill used to advance student understanding was to provide counter-examples to challenge student conceptions. His use of counter examples was similar to the idea of generative alternatives proposed by Rasmussen and Marrongelle (2006). However, the primary difference between Bill's use of counter-examples and generative alternatives is that generative alternatives typically serve to establish norms or promote student justification, whereas Bill typically used counter-examples to highlight deficiencies in student thinking. In addition, generative alternatives are a written record of some form, whereas Bill's counter-examples were purely visual or verbal.

In the following excerpt, Bill was leading a whole class discussion as the class worked to define “rigid motion”, a concept that was crucial to the class’s definition of symmetry[7].

Bill: If a rectangle is my figure, is that a rigid motion? [Bill rotates the rectangular eraser 90°.]
Class: Yes.
Bill: And what makes that a rigid motion?
S1: It has a pivot point.
Bill: It has a pivot point. What else?
Ellie: It’s not changing the object itself.
Bill: It’s not changing the object itself. Is this a rigid motion? [Bill moves the eraser 12 inches up and to the left.]
Class: Yes.
Bill: Does it have a pivot point?
Class: No.
Bill: Ok, so a lot of the rigid motions we have been dealing with have pivot points, it doesn’t seem to be a requirement of a rigid motion.

In this interchange, Bill did not accept the definition provided by student 1. Instead, Bill provided a counter-example that did not fit student 1’s definition of a rigid motion, but did fit the class’s intuition about what a rigid motion should be. By using a counter-example, Bill was able to clearly illustrate to the class that first student’s definition of rigid motion was inadequate, and he directed the class toward the desired definition.

Selective Restating

In addition to providing counter-examples, another move that Bill used to guide classroom discourse down productive pathways was selective restating. Within discourse analysis, the term revoicing has been defined as “the reuttering of another person’s speech through repetition, expansion, rephrasing and reporting” (Forman, McComrick & Donato, 1998, p. 531). O’Connor and Michaels (1993) claim that teacher revoicing is a useful teaching tool with many purposes, including allowing students to claim (or disclaim) ownership of ideas, giving credit to student ideas while allowing teacher influence to create “warranted inferences,” and lending authority to hesitant voices. Our use of the term restating includes both this definition of revoicing as well as asking students to restate their own speech. Asking students to restate previous ideas does not allow the teacher to directly expand on the student idea, however it has many of the other benefits of revoicing, such as lending authority and ownership to student ideas while influencing the focus of classroom discourse. In addition, restating can promote student empowerment through direct contribution, and it allows students the opportunity to practice reformulating their contributions for the class.

The following transcript except took place during whole class discussion of the definition of rigid motion, just after the dialogue illustrating Bill’s use of counter-examples presented previously.

Bill: And so, Ellie, what again, was your way of expressing that, what a rigid motion is? I did this [moves eraser 12 inches up and to the left], and you had a great explanation just now.

[7] Earlier in this class Bill had defined a symmetry as a rigid motion, that when applied to a figure, results in the figure landing on itself.

Ellie: Something that does not affect the object itself, from changing.

This conversation provides an example of asking a student to restate a previous idea. Bill selected a student contribution that he recognized as mathematically valuable, and he redirected the class's attention to that contribution. In this case, Bill did not question Ellie's definition further or add additional commentary. He accepted this idea and moves on to another topic.

Appealing to Definitions

Yet another pedagogical move that Bill made use of in this task was appealing to definitions. Previous to the conversation in the following transcript, the class had spent almost half of the class period reviewing the definition of symmetry. In this exchange, the students were asked to think about whether a 360° rotation was a symmetry or not. In the following passage, Bill engages the class in conversation.

> Bill: Let's see how people are landing on this. Um, how many people say a 360° rotation is a symmetry. [Many Students raise hands] How many people say that it is not? [Only one student raises a hand.] Let's hear from the not. Why isn't it a symmetry?
>
> S1: Uh, I just think it is kind of arbitrary. Because, like, if, if you want to be able to, like, count it, how are you supposed to count something if you can just, like, make it go as many times as you want it to.
>
> Bill: Well, the question we are asking is, does it meet the definition of symmetry? Is a 360° rotation a rigid motion?
>
> S1: Yeah.
>
> Bill: When you apply it to the figure, does it result in the figure landing on itself.
>
> S1: Yeah.
>
> Bill: So is it a symmetry?
>
> S1: Yeah.

The dissenting student in this passage partially articulated one of the dilemmas associated with equivalence. He stated that, "It is kind of arbitrary… [because you can] make it go as many times as you want it to." The difficulty that the student had might be solved by introducing the idea of equivalence classes; each of the rotations that are multiples of 360° fall into the same equivalence class and are thus not thought of as contributing to the total number of symmetries. Instead of building upon the mathematical insights of this student and opening up a discussion about equivalence, one of the goals of this task, Bill instead chose to appeal directly to the classroom's accepted definition of symmetry to convince the student that a 360° rotation is a symmetry. Bill's appeal to definitions directed the mathematical trajectory of the class away from the mathematically significant conflict experienced by one student and toward the classroom's accepted definition of symmetry.

Sequencing of Student Sharing

Organizing student contributions into logical order is one technique that teachers use to guide the flow of classroom discourse. Bill evidenced intentionality in the sequencing of the group presentations. In one instance during the small group time, Bill visited each of the groups and evaluated the maturity of each group's definition. He then selected the group that had the least complete definition to share first.

> Bill: This might be a good time just to share, kind of, where we are in the process for each group. And I want to go ahead and start with you guys because *I know that you are kind of still in the middle- more in the process of being formed*, but you have some of the idea. Can one or two or all 4 of you kind of articulate where you are with this?

The final group that Bill called on was the group that had the most complete and most formal definition. Bill had this group write their definition on the board and used it as a launching point for the next discussion.

Although Bill did make use of intentional sequencing to aid in the trajectory of lesson development, Bill's sequencing strategy is somewhat incomplete. Bill mentioned that he did not know the status of the second group's definition. Thus, although Bill indicated an awareness of intentional sequencing of some of the group presentations, his implementation did not take into account the state of one of the groups. Regardless, we see that in this case Bill was able to guide the trajectory of the lesson toward his intended goal even with this incomplete knowledge.

Conclusions

This work serves as a preliminary investigation into Bill's existent teaching techniques in his first implementation of an inquiry-oriented task. We found that he utilized a number of moves, including zooming out, illustrative real world examples, counter-examples, selective restating, appealing to definitions, and intentional sequencing of student sharing. Understanding the moves Bill utilized helps to paint a picture of some of the challenges he faced when opening his class to student-focused instruction and forms a basis upon which further professional development can occur. By identifying these teacher moves, it is indicated that mathematics instructors, even ones who primarily engage in teacher-center teaching, have techniques that they can draw on as they enact inquiry-oriented curriculum materials. Identifying such techniques can serve as a starting point for understanding how to support college-level teachers in changing their teaching practices. For instance, given the moves we identified, we can assume that Bill saw benefit in restating student contributions and utilizing examples (both counter and vivid). Therefore, motivating these tools may not need to be a focus of professional development. Instead, professional development could be designed to support Bill to shift the responsibility of restating and introducing examples to the students.

This study is just a starting point in an investigation of pedagogical techniques for student-centered teaching at the community college level. From this information, other important questions emerge, such as whether Bill will continue to use these moves through the semester, or whether the moves are modified or augmented through his experience with this curriculum. It would also be interesting to track Bill's pedagogy in courses outside of this class in order to study the influence of this inventive curriculum on Bill's pedagogy. Another possible goal for the extension of this research would the construction of a hypothetical learning trajectory (Freudenthal, 1991) for how lecture-based teachers learn how to implement inquiry-based curriculum. A more general hypothetical learning trajectory for teachers could be informative for professional development programs that support teachers in transition to student-oriented instruction.

References

Ball, D. (1993). With an Eye on the Mathematical Horizon: Dilemmas of Teaching Elementary School Mathematics. *The Elementary School Journal*, *93*(4), 373-397.

Boaler, J. (1998). Open and closed mathematics: Student experiences and understandings. *Journal for Research in Mathematics Education*, *29*(1), 41-62.

Boaler, J. (2006). "Opening our ideas": How a detracked mathematics approach promoted respect, responsibility, and high achievement. *Theory into Practice*, *45*(1), 40-46.

Cohen, D. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, *12*(3), 311-329.

Forman, E. A., McCormick, D. E., & Donato, R. (1998). Learning what counts as a mathematical explanation. Linguistics and Education, 9(4), 313–339.

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

Larsen, S. (2009). Reinventing the concepts of groups and isomorphisms: The case of Jessica and Sandra. *Journal of Mathematical Behavior, 28*(2-3), 119-137.

Larsen, S., & Bartlo, J. (2009). The Role of Task in Promoting Discourse Supporting Mathematical Learning. In L. Knott (Ed.), *The Role of Mathematics Discourse in Producing Leaders of Discourse* (pp. 77-98): Information Age Publishing

Larsen, S., Johnson, E., Rutherofrd, F., & Bartlo, J. (2009). *A Local Instructional Thoery for the Guided Reinvention of the Quotient Group Concept* Paper presented at the Twelfth Special Interest Group for the Mathematical Association of America on Research in Undergraduate Mathematics Education Raleigh, North Carolina

Larsen, S., Johnson, E., & Scholl, T. (2011). *Putting research to work: Web-based instructor materials for an inquiry oriented abstract algebra curriculum* Paper presented at the Fourteenth Special Interest Group for the Mathematical Association of America on Research in Undergraduate Mathematics Education

Lesh, R., & Lehrer, R. (2000). Iterative Refinement Cycles for Videotape Analysis of Conceptual Change. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 992). Mahwah, NJ: Lawrence Erlbaum Associates.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for school mathematics*. Retrieved from http://www.nctm.org/fullstandards/document/prepost/preface.asp

O'Connor, M., & Michaels, S. (1993). Aligning Academic Task and Participation Status through Revoicing: An analysis of a classroom discourse strategy. *Anthropology and Education Quarterly*, *24*(4), 318-335.

Phillippe, K. A., & Sullivan, L. G. (2006). *National Profile of Community Colleges: Trends & Statistics* (4th ed.). Washington DC: American Association of Community Colleges.

Rasmussen, C., & Marrongelle, K. (2006). Pedagogical Content Tools: Integrating Student Reasoning and Mathematics in Instruction *Journal for Research in Mathematics Education, 37*(5), 388-420.

Smith, J. (1996). Efficacy and teaching mathematics by telling: A challenge for reform. *Journal for Research in Mathematics Education*, *27*(4), 387-402.

Speer, N. M., & Wagner, J. F. (2009). Knowledge Needed by a Teacher to Provide Analytic Scaffolding During Undergraduate Mathematics Classroom Discussions. *Journal for Research in Mathematics Education, 40*(5), 530-562.

Wood, T. (2001). Teaching differently: Creating Opportunities for Learning Mathematics. *Theory into Practice*, *40*(2), 110-117.

HOW DO MATHEMATICIANS MAKE SENSE OF DEFINITIONS?

Margaret T. Kinzel
Boise State University
mkinzel@boisestate.edu

Laurie O. Cavey
Boise State University
lauriecavey@boisestate.edu

Sharon B. Walen
Boise State University
swalen@boisestate.edu

Kathleen L. Rohrig
Boise State University
KathleenRohrig@boisestate.edu

The purpose of this paper is to share preliminary results from a pilot study on mathematical definitions. Interviews with university mathematicians were designed to gain insight into mathematicians' processes for developing understanding of new definitions. We asked the participants to talk about what helps them understand a new definition and how they support students' understanding of definitions. We also observed them while they engaged in a definition task. Analysis revealed a noticeable difference in the emphasis on examples between what the participants described that they do and what they actually did while working on the definition task. We hypothesize that mathematicians' processes for making sense of a definition necessarily involve considering the definition's usefulness within a particular mathematical setting. Furthermore, these data indicate that mathematicians see examples as a multi-faceted, but not comprehensive, tool for understanding definitions.

Key words: definitions, advanced mathematical thinking, mathematicians, examples

Understanding mathematical definitions is essential for advanced mathematical thinking. However, the evidence suggests that even successful university students with substantial mathematics backgrounds think about and work with definitions in ways that are strikingly different than the practices of mathematicians (cf. Edwards & Ward, 2004; Tall & Vinner, 1981). Research has demonstrated that students prefer working with their own ideas about a concept (i.e. concept image) rather than a formal definition. Concept image is so powerful that it can be difficult to support students in developing an understanding of definitions that is more consistent with the mathematician's perspective. In their chapter on advanced mathematical thinking, Harel, Selden, and Selden (2006) made a case for mathematical definitions research that involves comparing the activity of students with the practice of mathematicians. In this paper we focus our attention on the practices of mathematicians. By doing so, we hope to learn more about what mathematicians actually do when they encounter a new definition, which in turn should inform the ways we support student learning.

Mathematicians encounter definitions in their work in a variety of ways. Definitions are part of the courses mathematicians teach, they are proposed by other mathematicians, and they are sometimes developed as part of the mathematician's own research. In instructional settings, mathematicians must decide how to present definitions to students. In the context of their mathematical work, mathematicians must judge the clarity and appropriateness of stated definitions. When preparing to propose a new definition, mathematicians must also consider presentation, clarity, and usefulness. Each of these settings requires some level of making sense of a given definition within a mathematical setting. We set out to create an interview context in which aspects of this activity were brought out and thus became accessible for analysis.

Specifically, we were interested in gaining insight into the processes mathematicians use to understand a new definition.

We situate our current work within existing literature on *thinking about* definitions while taking a grounded theory approach in our analysis. The ideas we share here contributed to our initial design and overall research question, but we did not attempt to apply an existing framework to these current data. Edwards (1997) drew attention to the lexical nature of mathematical definitions, that the formal statement of a mathematical definition places a concept within a particular class while also distinguishing it from other members of that class (e.g., *continuous function* as a particular type of *function*). Understanding the nature of mathematical definitions is necessary to support advanced mathematical thinking, as can be seen in the practice of mathematicians as summarized by Harel, Selden, and Selden (2006). In particular, their summary includes a list of features of definitions valued by mathematicians. That such a list exists indicates that mathematicians reflect on *definition* as a concept and on its role within the work of mathematics. The practice of mathematicians has been studied in terms of their use of writing, problem solving, and proving (e.g. Weber, 2008). While definitions play a role in each of these activities, the practice of mathematicians centered on encountering a *new* definition has not yet been characterized. Work has been done to investigate students' processes when encountering new definitions (cf. Zaslavsky & Shir, 2005; Zandieh & Rasmussen, 2010). Engaging students in the activities of evaluating statements of definitions and creating definitions seems to increase students' awareness of features of mathematical definitions, perhaps bringing them closer to ways in which mathematicians view definitions. In particular, features that are addressed across students and mathematicians include the need for definitions to be unambiguous, logically equivalent to other definitions of the same concept, hierarchical, stated in a usable form, and should address the purpose for which they were invented (Harel, Selden, & Selden). Definitions make it possible to sort a collection of mathematical objects into two distinct categories, those that satisfy a definition and those that do not satisfy that definition. Activity involving a definition, whether it be evaluating, creating, or using it, necessarily includes exploring representative objects (examples). Michener (1978) outlined distinctions between the roles played by examples within the work of mathematicians. Watson and Mason have written extensively on the role of examples, in particular with respect to student-generated examples (cf. Watson & Mason, 2002), and the potential for developing mathematical thought. We anticipated that asking mathematicians to reflect on their processes of making sense of definitions would involve the use of examples, both for themselves and for their students.

Our work is also grounded in a theoretical perspective on mathematics and understanding (Schoenfeld, 1994; Skemp, 1976). Schoenfeld described *mathematics* as the products of the work of mathematicians, *learning* mathematics as finding out about those products, and *doing* mathematics as creating those products by oneself or with others (1994, p. 55-56). Here we frequently reference understanding and sense making. We see understanding much like it was described by Skemp in that it is comprised of procedural and relational aspects (1976). When students learn mathematics, they are able to replicate certain procedures and recall facts. When students do mathematics, their work becomes consistent with that of mathematicians and results in relational understanding; knowing what to do, why, and how a mathematical idea is related to other mathematical ideas. Sense making refers to the psychological and socio-cultural processes involved in doing mathematics. When mathematicians and students are making sense of a mathematical idea they are developing their understanding of that idea.

Methodology

We conducted interviews with eight mathematicians who were employed at a large university in the northwestern region of the United States. The mathematicians volunteered to participate in the study. Five of the interviewed mathematicians were currently engaged in mathematical research in the areas of applied mathematics, cryptology, geometric topology, and set theory. Three of the mathematicians worked in multiple areas. All of the mathematicians were responsible for teaching at both the undergraduate and graduate levels and had experience that ranged from five to twenty or more years. Two were currently involved in small research studies aimed at improving instruction. In this preliminary report, we share findings associated with interviews of five of these mathematicians (Adam, Sam, Greg, Sadie, and Marc), three of whom were actively involved in mathematical research (Adam, Sam, and Greg).

The interviews were structured to gather data on mathematicians' thoughts and actions. During the first part of the interview we asked these mathematicians, "What helps you understand a new definition?" and "How do you help students understand definitions?" This part of the interview was designed to encourage them to describe what they do. We used follow-up questions to elicit specific instances of what they did to understand a new definition and what they did to help students understand a definition.

During the second part of the interview, we invited the mathematicians to engage in an example-generation activity and a definition task. The example-generation activity was taken from Watson and Mason (2005, p. 22). We asked the mathematicians to work through this task and then comment on its usefulness for teaching and learning definitions. For the definition task, we asked the mathematicians to talk aloud about what they were thinking as they attempted to familiarize themselves with a definition of formal language. (See Figure 1.) They were made aware that additional information was available upon request, which included definitions of related terms, like *support* and *formal power series*, and an example of a formal language. In addition, we explained that the goal during this part of the interview was to understand their processes, not for them to achieve understanding per se.

Figure 1. Definition of formal language

A *formal language* is the support of a formal power series over X^* where X is an alphabet. (Salomaa & Soittola, 1978)

The interview data were analyzed using a grounded theory approach described by Glaser and Strauss (1967). The interviews were transcribed to support analysis. We focused our initial analysis on the first two interview questions ("What helps you make sense of a new definition?" and "How do you help students understand definitions?") and on the definition task. In analyzing responses to the questions, first the researchers individually identified key aspects within a participant's articulations. Then, these individual aspects were brought to the group for discussion and negotiation of emergent categories. For the definition task, we worked in pairs to create a characterization of the participant's work on the task. These characterizations were then shared within the group and common categories were identified. We then began the work of comparing these categories across the mathematicians' articulated processes for themselves, for their students, and what we observed in their work on the definition task.

Results

Our presentation of results follows the organization of the analysis; that is, focusing on the mathematicians' responses to the two interview questions and their work on the definition task. In each section, we share brief excerpts to illustrate the emerging categories. In their articulations of their own processes, mathematicians clearly identified the use of examples as prominent. We were struck by the consistency between the mathematicians' descriptions of their own processes and their descriptions of how they structure their work with students. We initially noticed a lack of consistency between these articulations and our analysis of their work on the definition task. In particular, the use of examples shifted noticeably from a prominent place in their descriptions to a much more subordinate place in their work.

What I say I do. In response to the first interview question, the mathematicians shared what helps them to understand a new definition. The immediate response for four of the five participants included a reference to examples. This is exemplified in a quote from Adam, "I immediately try to think of examples of what would satisfy the requirements of the definition and what not." Adam stated that he has a set of "standard examples," such as the real line or the set of integers under addition, and that he uses these examples as test cases when he encounters new definitions. This testing process usually leads him to wonder "what things don't have these properties" and contributes to his understanding of the key features of the definition. Sam referred to "specific concrete examples," but acknowledged that finding such examples becomes difficult with abstract objects. Greg and Sadie both cited using examples and added that drawing diagrams or pictures is also helpful.

We asked the mathematicians to recall and share, if possible, an instance where they needed to make sense of a new definition. Adam, Sam, and Greg provided instances from their current research work. Examples played a key role in each of these instances. Adam described encountering the definition of a particular type of bounded group. He stated that he checked his standard examples (e.g., the real line) to see if they had the specific property. He also read a related theorem and realized that he could use his simple examples to build complicated examples that would have the desired property. However, this led him to wonder if there were other examples that did not arise from his simple examples. Adam indicated that examining a range of examples was necessary to understand what the definition intended to "capture."

Sam shared an instance in which he and his colleagues encountered collections of sets that intersect in a particular way. They wanted "to be able to find a definable family of objects with the property. So, it is not enough to know that there is one, we want to be able to isolate it somehow." Through looking at many varied examples, he stated that they identified the necessary conditions for the particular idea. Similar to Adam's process of understanding a newly encountered definition, Sam's newly created definition needed to capture the essential features of the collected examples.

Greg's instance came from algebra in relation to his work in topology. He shared that there are "humungous algebraic definitions that involve algebras with certain structures." He explained that since his work is primarily in geometry, his first inclination is to draw pictures. "So, even though those are algebraic definitions, I can still do what I like to do, namely draw my pictures." He described the process of "transform(ing) the definition into something more visual" so that he can "sit down and try to work out from the example what is really the essential feature of the definition." Across these three instances shared by the mathematicians, examples were prominent in their sense-making processes.

While explicit references to examples were evident in the responses, other aspects of the responses provide insights into more general processes. Marc and Sadie referred to seeing the definition “in context” as part of their process of sense-making. Sadie also mentioned connecting the definition to “other words that I might already know.” Adam shared that part of his process may involve “looking things up in books.” Sam noted a useful question: “Is this some notion that will actually be useful to prove the kind of results you are interested in?” Responses also included the need to attend to notation and language. Adam noted that one should pay particular attention to any quantifiers involved in the stated definition. Marc stated that the “phrasing has to get parsed; torn apart.”

From these data, mathematicians’ descriptions of their sense-making processes place examples in a prominent role. In particular, examples allow them to determine the essential features of a definition and to explore the boundaries of the defined collection of objects. We also get a sense that their processes involve active and purposeful engagement over a period of time. Each mathematician indicated that developing understanding of a definition requires focused attention on different aspects of the definition. The words, notation, context, and how particular terms are related to what they already know are important considerations for these mathematicians. Examples seem to serve as a primary method for focusing on one or more aspects of a definition. Furthermore, we see that mathematicians do not expect to necessarily have all the answers per se. Consulting texts and other sources for examples and other information was an important part of their articulated processes.

What I say I do for students. In response to the second interview prompt, the mathematicians talked about what they do to help students understand new definitions. We were struck by the similarity between their descriptions of their own processes and what they stated that they do in their work with students. Adam and Marc both explicitly stated that they attempt to match their students’ experience with their own. Adam said, “I try to kind of simulate for them this process I usually go through when I try to understand something.” The other participants, while not necessarily explicitly stating this intent, described activities that paralleled their descriptions of their own processes. Within these descriptions, participants referred to connecting new ideas to familiar ones and to the use of pictures or diagrams to help with visualization. Adam and Sam discussed the use of questions to focus students’ attention on the relevant features of a definition.

Consistent with their responses on their own work, the participants referred to their use of examples to support students’ understanding of concepts. Adam described two instances of using examples to introduce definitions in a number theory course. To introduce the idea of *isomorphism* between groups, Adam described a process of beginning with two groups that seemed different but that have “identically the same structure.” He described using the examples to focus students’ attention on that common structure before “writ(ing) down the precise mathematical criterion for this concept.” Once the definition was stated, Adam described engaging his students in applying the statement to an example involving two groups that appear to be similar; his stated reason is that it is harder “to show that things are not.” The second instance he described also used examples, but in this case he chose to present the stated definition first. Adam recounted how he presented the definition of a *group* and then led his students in a discussion using familiar examples of sets and operations to “explore which of those very familiar operations are really group operations and which ones are not and why.” In both instances, Adam’s description focused on the use of examples to highlight essential features of the defined concept.

Sam described a similar process more generally, "…part of giving them lots of examples is to hopefully turn this new thing into something somewhat mundane." Greg discussed the need to choose the right example, not too trivial but not too complicated, in order to focus attention on the essential features. Sadie "on a regular basis" gives examples of where definitions do not apply: "give an example where the limit definition doesn't work so you don't have a limit or where the function is not continuous and what part of the definition goes wrong." Adam summarized his use of examples and non-examples by saying, "I think it's important that one doesn't just do the positive, you know, where you confirm that something has the property. But, you also investigate the negative to see that, well, it's not just that everything is like this." Sam also shared a similar observation. "So, definitions exclude things. So, a definition tells you, okay, this is what we are looking at. But it also should tell you this is what we are not looking at."

Other responses referred to connecting the new definition to words or contexts that participants thought should be familiar to students. Greg described using real-world analogies to motivate the introduction of mathematical concepts. Sadie cited the usefulness of pictures to connect to familiar ideas. Marc explicitly referred to connecting to familiar terms. These references seem to be concerned with building a web of connections between mathematical terms, visual representations, and notations.

In discussing his own process, Sam commented on knowing how a concept might be used as part of building an understanding of the concept, "…because these might be very abstract objects, but then it might be useful to see, in the abstract setting you are working with, how is this used." This notion is paralleled in his articulation of his work with students. He commented that seeing this mathematical purpose can be difficult for students. "It might be they just don't have enough of the mathematical background intellectually to show the usefulness of these notions. … Eventually, if you are lucky in the course you are teaching, you get to some actual situations in which the notion, the new notion is actually useful." To address this, he gives students references to papers or books in which the notions are used. Sam stated that seeing the concept used for a mathematical purpose can support students' understanding of the idea, "Sometimes you are not going to, to gain any understanding of the definition from a specific example, or from a specific diagram, but only by really using it."

The mathematicians' articulations of their work with students clearly paralleled their descriptions of their own processes. Examples occupied a central role in how they described what they did to support students' understanding of definitions. The use of examples was supported by explicit connections to familiar ideas or images and by showing how a new idea might be used.

What I did. For the definition task we presented the mathematicians with a formal definition (Figure 1) from theoretical computer science with which we anticipated they would have limited familiarity. Our goal was to observe, in some sense, their initial sense-making processes. It was made clear in the interview that achieving understanding of this particular definition was not the goal; rather, we were interested in their thinking and in their reflections on their processes. Participants were aware that additional information related to the definition was available and would be provided upon request. This information is provided in the appendix for reference. An example was part of this additional information. However, not all participants knew that an example was available. Our analysis began with trying to capture the nature of the processes for each participant. We then looked across participants for common categories. Two broad categories emerged, (1) engaging in a decoding process and (2) connecting to familiar ideas or contexts. Given the prominence of examples in the first part of the interview, we also looked for references to examples in general or the use of the provided example.

The mathematicians began with a decoding process that involved obtaining additional information in the form of the supporting definitions. The participants differed in their apparent familiarity with terms within the initial definition and within the supporting definitions. Upon reading the initial definition, Adam offered that "this X is probably a set of things." He was given the definition for X^* and was able to confirm "so this is the set of all finite sequences." Sam's initial process was similar: "So, I think the first thing I will ask is what the definition of support is." After reading the definition for support, he stated, "I guess I immediately begin making some guesses as to what the other words mean." In some cases the decoding process focused on particular notations; Marc's first reaction to the initial definition statement was to note an unfortunate typesetting difference between the two X's on the page. While the typesetting was an error on our part, it revealed an aspect of the decoding process, that of understanding the use of notation and what it may imply about the concept involved.

The decoding process seemed to serve the purpose of identifying components that could be connected to familiar ideas or contexts. In some cases, this focused on particular uses of notation and what that might indicate about the intended context. For example, two of the supporting definitions included "   ", which seemed to lead Marc to question whether this referred to an interval or an ordered pair. In other cases, the supporting information allowed the mathematician to test connections to familiar contexts. Sam stated he was not seeing the use of the familiar *power series* in this context: "Okay, so I guess the one word that is not coming is what it means that a formal power series is over some set. So this is the word that I don't know how to apply in this context." For Adam, the supplied definition of *monoid* verified his conjecture, "Associative and an identity, okay. *Interviewer: Is that what you thought?* Yeah, and that one is okay." While some participants noted connections to familiar ideas and notation, none felt they achieved understanding of the definition. Adam stated, "So, at this point I'm not satisfied that I understand the definition. I will feel more comfortable when I'm at the point where I can actually reproduce these various things."

Given the mathematicians' emphasis on examples in the first part of the interview, we were initially surprised that references to examples did not come up earlier in their work on the definition task. Not all participants knew that an example was one of the pieces of available information; nevertheless, none of the participants asked for an example nor immediately introduced the need for one. For Adam and Sam, the example was offered when they seemed to reach a point of needing additional information that was not available:

> Adam: I am still confused about what is the semiring situation. And what is the zero. I can guess the zero is going to be the empty word, but what are the operations here? There's supposed to be two operations.
>
> Sam: I don't know when it says that this is a formal power series over X^*, then it means that X^* is the domain of the power series. Is that what it is saying?

When the available example was provided, Adam and Sam seemed to connect aspects of the example to familiar notions they had already identified, but having the example did not appear to immediately clear up their remaining questions. Adam said he would "try to think of things I would have thought of as a formal language and see if I can fit it into this." Sam stated, "this is much more general than I thought" and went on to say that he would need to see "what kind of operations one is interested in doing with language" in order to justify the level of generality.

The other participants' reactions to the provided example varied. Greg felt that he had achieved some level of understanding of the components of the definition and "now would have to do an example." In particular, he would need his example to help determine the importance of the semiring, a component of the definition he had not yet clarified. Sadie, after having asked for supporting information, noted that the interviewer still had an additional piece of paper. When presented with the statement of the example, she began a decoding process in a way that was similar to her work on the definition statement. When offered the example, Marc reflected on his processes: "That might be useful. And pretty clearly, my initial tendency was not in that direction. Certainly that's one way to approach a definition, instead of trying to understand its component parts." Adam, Sam, and Marc attempted to match the example with aspects of the definition. At the end of the interview the mathematicians stated that they would need additional information or to see how the definition would be used to feel more comfortable with the ideas.

While conducting the interviews, and in our initial analysis, we were struck by the consistency between what the mathematicians said they did for themselves and what they said they attempted to do for students and the apparent lack of consistency between what they said they do and with their work on the definition task. That is, the use of examples was a clear theme in what they described about their own mathematical work and their work with students. The mathematicians were articulate about the role that examples can play in developing understanding. However, when presented with a unfamiliar definition, the possibility of using an example to make sense of the concept was not part of their initial processes. We propose two themes from the data to explain our observations.

Themes

Mathematicians' processes for making sense of a definition necessarily involve considering the usefulness of the definition within a particular mathematical setting. Placing a definition within a setting involves a progression of previous definitions, notations, and examples. When presented with the unfamiliar definition, participants began by sorting through the specific terms and notations within the statement. This involved requesting and receiving various supporting definitions. Participants also made references to contexts with which they were familiar that contained terms or symbols used within the definition. In particular, participants questioned things such as why some terms were presented in a specific way, or whether the definition needed to be as general as it appeared to be. None of the participants asked to see an example as part of these initial processes. We see this as the mathematicians needing to situate the statement of the definition clearly within a mathematical setting to judge the value or usefulness of the definition.

Mathematicians see examples as a multi-faceted, but not comprehensive, tool for understanding definitions. In discussing both their own work and their work with students, participants spoke about examples as key to understanding. In particular, they noted that examples should be chosen carefully so that they serve to draw attention to important aspects. In making sense of definitions, the participants said they use examples to confirm their understanding, often choosing "messy" examples to be sure they had not introduced any inappropriate assumptions. Creating or considering non-examples was considered an essential component of understanding as articulated by Sam "…the only way to get there is look at concrete examples and look at concrete non-examples." The use of examples and non-examples seemed critical for the mathematicians' own understanding and how they went about supporting their students' understanding.

Our analysis related to this theme is ongoing. Some references to examples are more difficult to analyze than others; it is not always clear if when talking about "examples" the mathematicians referred to representative objects or to 'worked exercises' in the sense of using the new definition to accomplish some task. Nonetheless, we identified three potential facets of how the mathematicians described using examples as a tool for understanding: creating shared experiences, developing precision, and using test cases. Evidence for these facets come from participants' responses to the two interview questions and from their reflections throughout the interview. We share a few brief excerpts to illustrate our current thinking about these facets.

Examples can be used to create shared experiences working with a particular concept. Adam described how he used examples in his number theory class to explore the idea of *isomorphism*: "So we started with these two groups and we first played a little bit with this one, [I] tried to lead them to some of the essential properties." Sam stated a similar idea in reference to his own work: "Definitions come from isolating particular examples so that by the time you reach the definition, you have already acquired, hopefully, certain intuition about what this is." In addition, Greg described using a sequence of secant lines as a visualization to support students' ideas about tangent lines. From this, we infer that mathematicians sometimes use examples to familiarize their students with a particular concept. The intent seems to be to engage students in a process of exploration that is similar to what some of the participants described that they do for themselves. These shared experiences create opportunities for reflection upon the concept and upon aspects of the formal definition.

Examples can be used to develop precision. Participants repeatedly used words and phrases such as "capture," "know you have it," and "essential features" in their conversations with us. Such phrases seem connected to the lexical nature of mathematical definitions; that the formal definition clearly sorts mathematical objects according to precise criteria. The participants' references to using non-examples also seemed to be oriented towards developing precision. Several participants noted that it was not enough to know that an object was not a representative of a defined collection of objects. They expected their students (and themselves in their own work) to know which aspect of the definition was not satisfied.

Adam was explicit about how he uses test cases when he encounters a new definition. In both describing his research work and his instruction, he referred to thinking through a familiar set of objects and determining whether or not each object met the conditions of a given definition. This seemed to be a regular part of his initial process of exploring definitions. He noted that he acquired this "list of standard examples" by "playing with mathematics for years." That is, with experience, he developed both this particular process of exploration with test cases and expanded his repertoire of familiar objects.

In these interviews, mathematicians worked to place an unfamiliar definition within a particular mathematical setting. By way of contrast, the mathematical setting is already set in their general practice and in their instruction. When teaching, they attend to presenting ideas in a logical progression so that students have the necessary pieces to understand definitions as part of a particular mathematical setting. In their own work, they are familiar with current terms and notations within their field, so the setting and relationship between mathematical objects are understood. Within each setting, examples and non-examples serve to spotlight key features of the definition; using a non-example can help to clarify why certain aspects of the definition are needed; e.g. why does the function need to be defined at a point to be continuous at that point?

When we presented mathematicians with an unfamiliar definition, it was not clearly situated within a particular mathematical setting. They needed to understand the key components of the statement to establish the setting before they could use examples as a tool for understanding. That is, examples do not "carry" the definition entirely. Marc summarized this for us, "It's hard to ask for [an example] if you don't know what you're asking for."

Discussion

Our preliminary analysis indicates that mathematicians use a range of processes when making sense of a definition and that these processes are consistent with what they attempt to do in support of students' understanding of mathematical definitions. While the use of examples seems to play a prominent role in this work, examples are used in various ways and do not capture the full range of these processes. We can see, for example, that mathematicians expect understanding of a definition to take time and require focused attention on different aspects of the definition. Mathematicians use examples to explore the boundaries of a definition, establish key features of a definition, and as test cases. They ask questions to focus either their attention or their students' on a key feature or application of a definition. They consult resources when their questions go unanswered. Furthermore, we see the importance of knowing how the definition will be used within a mathematical setting. The participants described how they establish the purpose and context of a definition in their teaching by creating shared experiences working with examples. These examples are chosen to help students gain familiarity with new ideas in relation to what they already know. They also engage students in using definitions. In this way, the mathematicians attempt to replicate the processes for understanding that are an integral part of their own mathematical work.

The participants may not have been aware of the role of purpose when describing their processes of understanding. This is likely due to the fact that their work is usually accomplished within a very well-defined context. Our choice to present a formal definition outside familiar contexts for our participants allowed this aspect to become apparent. The juxtaposition of the mathematicians' articulated process with our observations of their actual process (albeit in an artificial setting) sets the stage for deeper analysis. Within a well-defined context, the role of examples may become more prominent, allowing that to be the focus of one's articulated process. With this observation in mind, we can now return to the mathematicians' descriptions to clarify references to purpose and context. Among the valued features of definitions (Harel, Selden, & Selden, 2006) are that definitions should address the purpose for which they were invented and that definitions should be hierarchical. The participants may not have made explicit references to these features; however, attention to such features may be implicit within their descriptions.

Since our observations are based primarily on what the participants described as their processes that necessarily places limits on any general conclusions we might make. Moreover, the participants' descriptions were somewhat varied and were influenced by what they were ready and willing to share with us. Although the descriptions of their processes were varied, we observed significant themes across the aspects of their work that were common to the participants. Our preliminary analysis of these data provide a broad range of snapshots into mathematicians' sense-making processes related to understanding definitions. The structure of the interview brought an implicit aspect of mathematicians' processes into focus, allowing for a more comprehensive characterization.

Future Work

As a preliminary report, we conclude with future tasks and directions. Analysis of the interviews with the other three participants is currently underway. We have not yet analyzed the interview data on the example generation task. Participants worked through and reflected on the task. These data may provide further insights into the mathematicians' perception of the role of examples, both for themselves and for their students.

Another task ahead of us is to go back to the data and verify (as best we can) exactly what the mathematicians were referencing when they used the word "example." We want to clarify when they might have been thinking of a 'worked exercise' as opposed to representative object. We anticipate using the work of Michener (1978) and Watson and Mason (2002; 2005) to assist us as we analyze these references to examples.

In addition, we conducted a parallel set of interviews with students. Students from a variety of mathematics courses were asked to describe how they make sense of definitions and to provide a description of an instance where they made sense of a definition. We plan a similar grounded-theory approach to the analysis of these data in an effort to reveal common categories that can be used to compare or contrast with those found in the mathematicians' data.

Finally, we note that context became a salient aspect within the interviews with the mathematicians. In some sense, the presented definition may have been too removed from familiar contexts for us to clearly see the participants' sense-making processes. Using a definition more closely related to their mathematical experiences may enhance our understanding of the mathematicians' approaches to making sense of definitions.

References

Edwards, B. S. (1997). Undergraduate mathematics majors' understanding and use of formal definitions in real analysis. (Unpublished doctoral dissertation.) Pennsylvania State University, State College, PA.

Edwards, B. S. & Ward, M. B. (2004). Surprises from mathematics education research: Student (mis)use of mathematical definitions. *The American Mathematical Monthly, 111*(5), 411-424.

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

Harel, G., Selden, A., & Selden, J. (2006). Advanced mathematical thinking. In A. Gutierrez & P. Boero (Eds.), *Handbook of research on the psychology mathematics education: Past, present, future* (pp. 147-172). Rotterdam, The Netherlands: Sense Publishers.

Michener, E. R. (1978). Understanding understanding mathematics. *Cognitive Science, 2*, 361-383.

Salomaa, A. & Soittola, M. (1978). *Automata-Theoretic Aspects of Formal Power Series*. Springer-Verlag: New York.

Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In A. H. Schoenfeld (Ed.), *Mathematical Thinking and Problem Solving* (53-70). Lawrence Erlbaum Associates: Hillsdale, New Jersey.

Skemp, R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77*, 20-26.

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.

Watson, A. & Mason, J. (2002). Student-generated examples in the learning of mathematics. *Canadian Journal of Science, Mathematics and Technology Education, 2(2)*, 237-249.

Watson, A. & Mason, J. (2005). *Mathematics as a Constructive Activity: Learners Generating Examples*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39(4)*, 431-459.

Zandieh, M. and 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.

Zaslavsky, O. and Shir, K. (2005). Students' conception of a mathematical definition. *Journal for Research in Mathematics Education, 36*(4), 317-346.

Appendix

Monoid: A *monoid* consists of a set *M*, an associative binary operation * on *M* and an *identity* element 1 such that 1*a = a*1=a for every a.

Semiring: A *semiring* is a set *A* together with two binary operations + and * and two constant elements 0 and 1 such that

1. <*A*,+,0> is a commutative monoid,
2. <*A*,*,1> is a monoid,
3. The distributive laws a*(b+c) =a*b + a*c and (a+b)*c = a*c + b*c hold, and
4. 0*a = a*0 =0 for every a.

Formal power series: Let *M* be a monoid and *A* a semiring. Mappings *r* of *M* into *A* are called *formal power series*. The values of *r* are denoted by where $w \in M$. We denote *r* by

$$r = \sum_{w \in M} (r, w)w .$$

Support: $\{w \mid (r,w) \neq 0\}$ is called the *support* of and denoted by *supp(R)*.

Alphabet: An *alphabet X* is a finite nonempty set.

$X^* = \bigcup_{n=0}^{\infty} X^n$ where $X^0 = \{\lambda\}$, $X^2 = \{xy \mid x, y \in X\}$, and for n>1. Note λ is called the *empty word*.

Example: Let $X = \{a,b\}$, then $L = \{a^n b^n \mid n \in N\}$ is the support of .

SPANNING SET AND SPAN: AN ANALYSIS OF THE MENTAL CONSTRUCTIONS OF UNDERGRADUATE STUDENTS

Darly Kú
CINVESTAV-IPN
darlyku@cinvestav.mx

Asuman Oktaç
CINVESTAV-IPN
oktac@cinvestav.mx

María Trigueros
ITAM
trigue@cinvestav.mx

Abstract

We present a genetic decomposition using APOS Theory, about the way in which students can construct the concepts of spanning set and span in Linear Algebra. We also present empirical data coming from interviews made with 11 university students who had completed a course on Analytic Geometry which included an introduction to Linear Algebra. We report on our observations and suggest possible modifications to our initial theoretical analysis.

Key words: Spanning set, APOS Theory

Introduction

It has been reported in different research studies that learning Linear Algebra is difficult for university students and that many of these problems are related to the difficulties students face when learning abstract concepts (Dorier, 1997). Several studies have been carried out in different countries, and some authors have developed didactical proposals to teach specific concepts, such as vector space, linear transformations, and basis, or designed methodologies in order to help students to overcome the detected obstacles (Sierpinska, 1994; Alves Días& Artigue 1995; Harel, 1997).

In an earlier study about the construction of the concept of basis in Linear Algebra (Kú, Trigueros & Oktaç, 2008) we observed the difficulties that students have with the concept of spanning set and the coordination of the underlying process with the process related to linear independence. These difficulties seemed to interfere in a serious manner with the construction of an object conception of basis of a vector space. As a result we decided to carry out research in order to look at these concepts separately, so that we could offer an explanation about the construction of each concept and related problems.

Some literature published previously touch certain issues related to the learning of spanning sets focusing on task design, cognitive difficulties and suggestions for teaching (Nardi, 1997; Ball et al., 1998; Dorier et al., 2000; Rogalski, 2000). What we are interested in with this research is to offer a viable path that students may follow in order to construct this concept as well as explaining the nature of related difficulties while learning it. After completing our analysis of the empirical data, we also hope to make pedagogical suggestions.

Our research questions in this study are:

Do the constructions modeled in a genetic decomposition for the concepts of spanning sets and span useful in describing students' ways of dealing with problems related to these concepts?

What constructions can be associated to students' difficulties?

How can this information be used to make suggestions for the teaching of these concepts?

APOS Theory

APOS theory has been used successfully in explaining the construction of several concepts in undergraduate mathematics curriculum, and in designing theory based didactic approaches to teach them (Dubinsky, 1996). It has been proved to be effective for both purposes by means of studies concentrating in many different mathematical areas, such as Abstract Algebra, Differential and Integral Calculus, Statistics, and Discrete Mathematics (Dubinsky & McDonald, 2001). Its effectiveness is related to the possibility to explain certain processes that take place when learning advanced mathematics topics using the conceptual tools developed in the theory. Its use to study the construction of Linear Algebra concepts is more recent but results obtained so far are promising (Trigueros & Oktaç, 2005a; Roa-Fuentes & Oktaç, 2010; Parraguez & Oktaç; 2010; Trigueros, Oktaç & Manzanero, 2007). We continue with this line of research and use APOS theory to study the mental constructions and mechanisms involved in the learning of spanning sets.

APOS theory (Action, Process, Object, Schema) is an adaptation of Piaget's epistemological ideas to the learning of advanced mathematics at the university level. According to APOS theory an individual's mathematical knowledge and its development can be defined as:

"An individual's mathematical knowledge is her or his tendency to respond to perceived mathematical problem situations by reflecting on problems and their solutions in a social context and by constructing or reconstructing mathematical actions, processes and objects and organizing these in schemas to use in dealing with the situations" (Asiala, Brown, DeVries, Dubinsky, Mathews & Thomas, 1996).).

In this definition the main elements that enable a researcher to discern the way in which a student understands a mathematical concept, and that constitute the fundamental constructs of APOS theory are the mental structures called action, process, object and schema.

We say that students have an action conception of a mathematical concept if they perform transformations on objects in the form of step by step calculations or if they rely on memorized facts. According to Asiala et al. (1996) "an individual whose understanding of a transformation is limited to an action conception can carry out the transformation only by reacting to external cues that give precise details on what steps to take". Even if an action conception constitutes a limited form of understanding, the construction of action conceptions is crucial as a starting point to construct a concept.

After repeating actions on objects and reflecting upon these actions, students interiorize them into a process. Students who show a process conception are able to think about the transformations and describe them without a need to perform each step explicitly.

When students reflect on the processes and are able to think of them as a whole, we say that they have encapsulated the process of applying a transformation into an object. This implies that they are able to apply actions on the newly constructed objects. Students who show an object conception of a concept are also able to de-encapsulate an object into the process from which it originated.

Actions, processes, objects and other schema, and connections between them form what is called a schema in APOS theory. Therefore it can be said that a schema for a mathematical concept is a coherent collection of actions, processes, objects, and other schema that are related as a structure in an individual's mind and that can be used in a problematic situation related to a particular mathematical topic or concept. The coherence of the schema refers to the ability of the individual to decide whether it is possible to work on a mathematical situation using that schema (Dubinsky & Mc Donald, 2001).

According to APOS theory, the mechanism used in transitions from a type of conception to another is abstractive reflection. Trigueros (2005b) mentions that this mechanism is activated through the physical or mental actions that an individual makes on a knowledge object.

Transitions from action to process or from process to object do not occur in a linear way. Students can show action conceptions for some aspects of a given concept they are learning, and show, for example, a process conception of other aspects of the same concept. It can be even difficult to decide the kind of conception a student shows because it is possible to interpret her or his explanation for a specific problem where some elements can be regarded as actions or processes and others as processes or objects. Also it is important to clarify that students' responses to problematic situations are necessarily related to the cognitive demand required by the problem. If the solution of a problem only requires actions, students will certainly use actions to work on it; which does not mean that they cannot show another kind of conception in another problem situation.

When using APOS theory it is necessary to develop an idealized and detailed description of the actions, processes, objects, schemas and their relationships occurring in the construction of a mathematical concept. This model is known as a genetic decomposition of the concept in question. The viability of a genetic decomposition can be tested empirically using students' work. The results of the analysis of the data obtained from this source are used to refine the genetic decomposition so that it gives a better description of the way students construct that concept. A genetic decomposition can also be used as a guide in the design of teaching materials. It is important to clarify that several different genetic decompositions can exist for the same mathematical concept; what is important, however, is for any genetic decomposition to describe what is observed in students' work.

Based on this conceptual framework the research reported in this paper followed these steps: First we developed a possible genetic decomposition for the concept of spanning sets, afterwards we designed a research instrument to probe students' mental constructions and to test the viability of our genetic decomposition. We are in the process of analyzing students' responses to the questions in the instrument to describe the mental constructions that students have made relative to the concept of spanning sets. Now we present the work involved in each of these steps, commenting on some of the results that we obtained.

Preliminary Genetic Decomposition for the concept of spanning set

According to the methodology linked to the APOS theory "research begins with a theoretical analysis modeling the epistemology of the concept in question: what it means to understand the concept and how that understanding can be constructed by a learner. This initial analysis, marking the researchers' entry into the cycle of components of the framework, is based primarily on the researchers' understanding of the concept in question and on their experiences as learners and teachers of the concept" (Asiala, et al. 1996). It is worth mentioning that in our genetic decomposition we include the constructions we consider necessary for the students to differentiate the meaning of the two concepts: spanning sets and spanned space.

Prerequisites

One possible set of concepts to start the construction of the notions of spanning set and spanned space are vector space, variable and solution set for a system of equations.

- Vector space concept is fundamental in the construction of several other linear algebra concepts, in particular those of spanning set and spanned set. We consider that students need to be able to recognize some familiar vector spaces that they have worked with, such as spaces with dimension 1, 2 or 3 and with defined elements (such as n-tuples). Students must also recognize that there are sets containing elements other than numbers that can be considered as vector spaces, for example, polynomials and matrices, and to be able to work with them.

- We consider that the solution set of a system of equations plays an important role in the interpretation of the meaning of spanning set and spanned space, so we consider that the students should demonstrate an object conception of this concept. According to Trigueros et al. (2007), students demonstrate an object conception of this concept if they can represent in parametrical form the solution set of a given system or describe its geometric representation accurately.

- Variable is another concept we consider important to start the process of construction of the above mentioned concepts. We consider that students need to work with variables as mathematical objects and so they need to understand variables as unknowns, general numbers, or parameters, and variables in functional relationships, and move flexibly between all these representations (Trigueros and Ursini, 2003).

Mental Constructions

Given a vector space V, a specific set S of vectors from V and a specific scalar field K, students need to perform actions on the vectors of S and the scalars. These actions consist of performing scalar products and sums of vectors in order to obtain a new vector of V. Coordination of these actions is interiorized into the process of construction of a new vector which is an element of the vector space, that is, into the process of construction of linear combinations. This process also implies that the student can verify if a given vector can be written as a linear combination of a given set of vectors. We consider this as a process conception of linear combinations.

Through actions or processes on a given set of vectors S, students can verify if there are scalars in K that can be used to express the elements of a new set of vectors T in the vector space V as a linear combination of S. This process is coordinated with the process of finding the solution set of the resulting system of equations taking into account the notion of variable. The result of this coordination is the process of finding a set of scalars. Through the action of forming the linear combination with these scalars and S, the student can verify that S generates T. This process is generalized to include different instantiations of S and T. When the set S can be considered as a whole, whenever it is needed, this process has been encapsulated into an object that we may call spanning set.

The object conception implies that students are able to explain that a given vector space can be generated by different spanning sets, that these sets do not necessarily have the same number of elements, or common elements, and that the number of elements in general is not the only determining factor if a set is a spanning set for a vector space.

By reversing the last process students can construct the spanned set of vectors. They can also perform on it the actions needed to verify that this set T is a vector space or a subspace of a given vector space V. This process can be generalized to determine if V can be formed by all the possible linear combinations of the set of vectors T. This generalization process is encapsulated into a new object that can be called a generated set, or spanned set, by the given set of vectors. These constructions must enable students to differentiate the concepts of spanned space and spanning set.

It is important to point out that we are not ignoring those concepts that are related to spanning sets, such as linear independence or dependence, basis, dimension, etc. We emphasize that our analysis is focused on how the construction of the spanning set can help students understand other concepts related to it, or if there are difficulties in the construction of this concept that act as obstacles when relating it with other Linear Algebra concepts.

We are also aware that the construction process can be started by the construction of the spanned (generated) set and followed by the construction of the spanning set. Experimental data would be needed to compare these different construction processes.

Based on this genetic decomposition we can identify some observable behaviors which we can link to the described mental constructions that students might display when solving problems related to the concepts of spanning set and spanned space.

We will consider that students have an Action conception of spanning set and spanned set if they show a process conception of linear combination, can use it to verify if there are elements of K that can be used to write a specific vector from V as a linear combination of the vectors of the set S. When verifying if the set spans the given vector space the student can do it only with specific vectors.

If the student is able to generalize the previous actions and running through the elements of S in her/his mind considers that every vector of V can be expressed as a linear combination of the elements of the set S, we will say that this student displays a process conception of spanning sets.

A student will display an object conception if she or he demonstrates that he or she can apply actions on the spanning set or the spanned set. This implies, among other things, that he or she considers that the spanning set is not unique, that different spanning sets do not need to have common elements, and that the number of elements of a set by itself is not a valid criterion to determine if the set is or not a spanning set. It is also important that she or he considers the fact that the generated space is the result of all the possible linear combinations of the spanning set, and that these two concepts are different.

Methodological aspects

We then designed an interview that consisted in 7 questions, in order to test the viability of our genetic decomposition. This instrument was applied to a group of 11 undergraduate students who were taking an analytic geometry course at a Mexican university. This course consists of an introduction to Linear Algebra with a strong emphasis on the geometric interpretation and properties of vectors in R^n, but it also covers vector spaces with matrices and polynomials as elements. In our design of the interview questions we took into account different aspects of a spanning set. We asked questions of the type whether a certain set spans a given vector space, but we also asked the construction type of questions, namely given a vector space identifying possible spanning sets for it. We also asked the students to compare the vector spaces generated by different spanning sets. By dealing with different aspects of the concept of spanning set in this manner, we hope to shed light on where the difficulties lie and verify the related mental constructions.

Interviews were tape-recorded and all the written information was kept as part of the data set. These interviews were analyzed according to our theoretical framework. The analysis focused on the identification of the mental constructions used by students while working with the problems and their comparison with those modeled in the preliminary Genetic Decomposition. We also focused on how students related different mathematical concepts.

Results and Discussion

In order to illustrate some difficulties related to the construction of the concept 'spanning set', in this section we give some examples from the interviews that we conducted, interpreting the data from the theoretical lens of APOS theory. In the extracts below students are presented by pseudonyms and 'I' stands for the interviewer. The two interview questions we selected for this purpose are numbered (3) and (5).

3. Let W be the subspace of R^3 that consists of all the points (x, y, z) such that $x+3y-4z=0$. Find a spanning set for W that contains two vectors. Can only one vector span W? Can three vectors span W? Justify your answer.

Question 3 has the purpose of identifying the strategies that students use, in determining a spanning set for a plane given by an equation. We also want to find out if students can make connections between the number of vectors in a set and its spanning properties.

We start with Javier, who writes the following after reading question (3) and comments on it:

W subespacio de $\mathbb{R}^3$ (x, y, z)
$x+3y-4z=0$ $(1, 3, -4)$

Javier: I think it can be written this way, the x, the value of y, the value of z, and I think I can put it like this (writes):

$$\begin{pmatrix} 1 \\ 3 \\ -4 \end{pmatrix}$$

Now I would have to think how to span it, and they are asking me to span it with two vectors, right? So I think it can be put as $x=4z-3y$… (He writes $\begin{pmatrix} 1 \\ 3 \\ -4 \end{pmatrix} = a\begin{pmatrix} \\ \end{pmatrix} + b\begin{pmatrix} \\ \end{pmatrix}$). This (he points to the empty parenthesis multiplied by a) times a and this (writing the other empty parenthesis multiplied by b) times b.

Javier: This one (referring to the first vector), I want to decompose it in two so that when we do it, after we multiply it by this (referring to a) and by this (referring to b) and then we add them, it gives me the subspace of R^3.

When the interviewer asks him how he can check if the set he proposes really spans W, Javier does the following calculation and declares that he's done.

$$a\begin{pmatrix} 4 \\ 0 \\ -4 \end{pmatrix} + b\begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 4a-3b \\ 3b \\ -4a \end{pmatrix} \quad \begin{matrix} b=y \\ a=z \end{matrix}$$

$$\begin{matrix} 4z-3y=x \\ 3y \\ -4z \end{matrix}$$

The interviewer asks again how he knows that the two vectors span W and he replies by saying that the two vectors are linearly independent. When the interviewer insists on the question of how to check the spanning of W by these two vectors, Javier says he doesn't know how to do it.

In the extracts shown above, we can observe that Javier chooses a vector with which to work, that is the normal vector of the given plane. Possibly the fact that a plane is determined by its normal vector (fact that is emphasized in the Analytic Geometry course he had been taking) makes him believe that it can be chosen as the representative of the given plane and that by working with it he would be working with the plane. Then he writes the normal vector as a linear combination of two other vectors. The answer given by Javier shows that there is a lack of coordination between the processes of belonging to a set and forming linear combinations, which is necessary in order to have a process conception of spanning sets. Javier, by only working with one of these processes, namely forming linear combinations, and not coordinating it with the other one, is not able to find a spanning set: the two vectors he proposes do not belong to the subspace in question. Afterwards, Javier comments on the possibility of spanning the given subspace by only one vector, or three vectors. However, obviously he is working with another subspace, the one generated by the normal vector.

> Javier: Then I will go to … if only one vector spans W, and I think yes, because since it is a subspace actually it is not the whole of R^3, and since it's only a… I think it's only a line, so I think it can be spanned by only one vector. And I think the third one is 'can three vectors span W?'. Yes, but they won't be linearly independent, since if they were linearly independent they would span all of R^3… I think.

Implicit in his answer above is the assumption that if a vector v_1 can be written as a linear combination of two other vectors v_2 and v_3, the subspaces generated by the two sets $\{v_1\}$ and $\{v_2, v_3\}$ are the same. Furthermore, it seems that Javier only makes a difference, in terms of dimension, between two kinds of subspaces: the vector space itself and all other proper subspaces, these latter ones all being put into the same category.

Then when he is asked to describe what it is that makes a spanning set, the following interchange occurs:

> Javier: By means of combinations of the elements it can generate a space.
>
> I: Ok, so how could you check whether that really is a spanning set for W, taking into account what you have just said?
>
> Javier: Well I am thinking in another characteristic because I see that these are linearly independent and combining them it comes out but there has to be another characteristic in order to determine if with that it will be a spanning set.
>
> I: So you make the linear combinations, and what does that imply?
>
> Javier: I make the linear combinations and it gives me the set.
>
> I: The set. Which set?
>
> Javier: The generated set.

One of the difficulties Javier is experiencing has to do with a confusion between two types of tasks: generating a subspace by using a set of given vectors, and given a subspace finding a possible set of spanning vectors. This and other data imply that a student might have different conceptions regarding the two concepts spanning set and generated subspace, and the related constructions may not be connected in the student's mind.

Now we pass to Oscar who makes use of the normal vector as well when he starts working on the problem, but uses a different strategy to proceed.

> Oscar: (writes)
>
> $$x+3y-4z=0$$
> $$N=(1,3,-4)$$
> $$n_\perp = (4,0,1) \quad (1,3,-4)\bullet(4,0,1)=0 \quad \text{and} \quad P\,(5,1,-2)$$
>
> So here what I did was, I have the equation of a plane in R^3, I find the normal (vector), then I find the perpendicular to the plane, well making sure it gives zero. When I have it I look for a point that passes through the plane and well with that I can express the equation of the plane (writes):
>
> $$\Pi = \{P|P= (5, 1,-2)+ t(4, 0,1)\}$$
>
> But I need another… a point of the plane, the normal, I need another one. With only one vector I cannot span W.
>
> Oscar: Because in the equation of the plane I have two parameters so well, I have any point and two parameters, right? Well two direction vectors and parameters in the reals. So... but… in order to span W which is an equation of the plane in R^3, I always need two vectors. So with one I would be spanning a line.
>
> I: Ok
>
> Oscar: And with three vectors I could be sp... if they are linearly independent I could be spanning R^3 and if they are not linearly independent no… Supposedly it spans a plane but if they are not linearly independent I cannot make sure that they span. But with two vectors I can span W.

At this point the interviewer goes back to the first problem and asks how he would obtain the spanning set that is being asked for. Oscar responds by saying that he doesn't remember. When he is asked what is required for obtaining a spanning set, Oscar gives the following answer:

> Oscar: Being linearly independent and belonging to the set and being linearly independent in order to span it.

Later he mentions that the set {(5, 1,-2), (4, 0,1)} could be a spanning set, but that he would have to verify linear independence. He writes a system of linear equations and by solving it he affirms that the set is linearly independent and hence a spanning set for W. Although the vector (5,1,-2) does not lie on W because of the negative sign, it can be observed that Oscar is beginning to coordinate the two processes that give rise to the process of spanning a subspace. However in his interview too we observe that he has different types of conceptions related to the two concepts 'spanning set' and 'spanned subspace'.

5. Let $v_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, v_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$ and let $H = \left\{ \begin{bmatrix}s\\s\\0\end{bmatrix} s \text{ in } R \right\}$. Therefore each vector of H is a linear

   combination of $\{v_1, v_2\}$, since $\begin{bmatrix}s\\s\\o\end{bmatrix} = s\begin{bmatrix}1\\0\\0\end{bmatrix} + s\begin{bmatrix}0\\1\\0\end{bmatrix}$

Answer the following question: Is $\{v_1, v_2\}$ a spanning set for H? Justify your answer.

With question (5) we emphasized explicitly the belonging of the vectors of a spanning set to the subspace that they generate.

Pedro gives the following response to this question:

> Pedro: They are giving me these two vectors and the set of elements. Well, the set will be all the elements that are in this form, where their first two elements are equal (s,s,0), so a linear combination of v_1 and v_2 will give you all of H, but not the combination of v_1 and v_2; that is there will be things that they generate that will not be in H. They will generate all of H, but also more things.
>
> I: So is it a spanning set for H or not?
>
> Pedro: No, it is not because they will generate even more. Yeah, I mean the linear combination of these two, I mean the span of v_1, v_2 will be greater than the set H.
>
> I: And what could be a spanning set for H?
>
> Pedro: A set… it could be for example the vector (1, 1,0), that is the set H is a set of dimension 1. You only need… you can only have one and that will span everything.

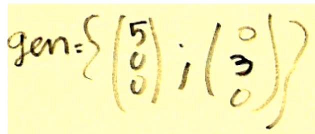
As the above extract shows, Pedro does not base himself only on linear combinations. Although he doesn't make explicit reference to belonging of the vectors to the subspace, he employs the idea that the linear combinations formed by the vectors of a spanning set should exactly generate the given subspace, and not more.

On the other hand, Carlos has difficulties in deciding whether the given set of vectors form a spanning set for H.

> Carlos: (writes)
>
> "Yes. All possible linear combinations of v_1 and v_2 span H. Neither v_1 nor v_2 can generate the third element, but H doesn't have it either, so it is not necessary."
>
> (then he explains)
>
> It's a spanning set for H because if we take all the possible linear combinations in the reals, then clearly we can see that it can be any number… well, any number in H and for example H doesn't have… it has a zero in the third element so, no, well…it would be, it's not needed and we see it here… I mean none of the two has it so… If H had another *s* here for example (he refers to the vector (s,s,0)), it wouldn't be a spanning set for H, we would need another vector, which had for example, I don't know. If it were linearly independent and if it had an element in the last position, but since these two don't have it, but H doesn't either, then H can be spanned by these two vectors.

Although Carlos mentions "all possible linear combinations", he doesn't make any reference to the belonging of the vectors v_1 and v_2 to H. The interview continues with the following question, in order to provoke a reflection in Carlos:

> I: Can you give another spanning set for H?
>
> Carlos: (writes)



> I: Let's see. Why is this a spanning set?
>
> Carlos: Because they are two linearly independent vectors and if we take any numb… I mean this is in reals, so if we take any number in H, well for example in H, I don't know, for it to be 1 and 1. We multiply this one by 1/5 added to this one multiplied by 1/3 and it spans H.

We can observe that Carlos still wants to obtain the vectors v_1 y v_2 in order to find a different spanning set. When the interviewer asks how he would find a spanning set for H if the question didn't give the set $\{v_1, v_2\}$, Carlos responds as follows:

> Carlos: If I didn't have this? Well, *s* is in the reals, so it could be any number. Well, it would be enough to take two vectors that, I mean, with which I can generate a real number in the first one and a real number in the second one and that would be enough.

As we can see Carlos keeps thinking about a similar set to the one given in the question in order to span H, and does not bring into the situation the necessity of the vectors belonging to H.

Conclusions

In this paper we propose a genetic decomposition that models how students might construct the concepts of spanning set and span. The empirical evidence presented illustrates certain difficulties related to this construction. Preliminary results show that in order to find a spanning set of a given vector space, a process conception of the related concept is needed, since the task requires the coordination of two underlying processes: belonging to a set and linear combinations. Students tend to ignore the belonging part when working on situations involving the concept of spanning set.

Students who have a process conception of the concept of spanning sets can relate it to other concepts such as linear independence and dependence, and dimension. We also observe that in general it is easier for students to decide whether a given set spans a given vector space than constructing a spanning set for a given vector space. These data inform our theoretical analysis and we propose some modifications in our genetic decomposition to reflect these results. On the one hand there should be an explicit mention of two processes that need to be coordinated as mentioned in the paragraph above, and on the other hand it should be emphasized that the two concepts spanning set and span may be evolving conceptions that involve different types of constructions.

Another observation is that even mathematically speaking the coordination of the two processes spanning and being spanned seems to be simultaneuous, cognitively speaking it may not be so.

Reference

Alves Dias, M. & Artigue, M. (1995). Articulation Problems Between Different Systems of Symbolic Representations in Linear Algebra. Proceedings of the 19th annual meeting of the international group for the Psychology of Mathematics Education, Recife, Brasil, vol. 2, 34—41.

Asiala, M., Brown, A., DeVries, D., Dubinsky, E., Mathews, D. & Thomas, K. (1996). A Framework for Research and Curriculum Development in Undergraduate Mathematics Education. Research in Collegiate Mathematics Education II, (6), 1-32.

Ball, G., Stephenson, B., Smith, G., Wood, L., Coupland, M. and Crawford, K. (1998). Creating a diversity of mathematical experiences for tertiary students. ,*International Journal of Mathematical Education in Science and Technology*, **29**(6), 827-841.

Dorier, J. L., Robert, A., Robinet, R. and Rogalski, M. (2000). The Obstacle of Formalism in Linear Algebra. A Variety of Studies from 1987 Until 1995. In J.-L. Dorier (ed.), On the Teaching of Linear Algebra. Dordrecht : Kluwer, pp. 85-124.

Dubinsky, E. (1996). Aplicación de la perspective piagetiana a la educación matemática universitaria . *Revista Educación Matemática*, 8(3), 24-41.

Dubinsky, E. & McDonald, M. (2001). APOS: A Constructivist Theory of Learning in Undergraduate Mathematics Education Research. In D. Holton, M. Artigue, U. Kirchgräber, J. Hillel, M. Niss & A. Schoenfeld (Eds.), *The teaching and Learning of Mathematics at University Level: An ICMI Study* (pp. 273-280), Kluwer Academic Publishers.

Harel, G. (1997). The linear algebra curriculum study group recommendations: Moving beyond concept definition. In Carlson D., Johnson, C, Lay, D., Porter, D., Watkins, A, & Watkins, W. (Eds.). Resources for Teaching Linear Algebra,. MAA Notes, Vol. 42, 107-126.

Kú, D., Trigueros, M. & Oktaç, A. (2008). Comprensión del concepto de base de un espacio vectorial desde el punto de vista de la teoría APOE. *Revista Educación Matemática*, *20* (2), 65-89.

Nardi, E. (1997). El encuentro del matemático principiante con la abstracción matemática: Una imagen conceptual de los conjuntos generadores en el análisis vectorial. *Educación Matemática,* **9**(1), 47-60.

Parraguez, M. & Oktaç, A. (2010). Construction of the Vector Space Concept from the Viewpoint of APOS Theory. Linear Algebra and its Applications, 432, 2112-2124.

Roa, S. & Oktaç, A. (2010). Construcción de una descomposición genética: Análisis teórico del concepto transformación lineal. *Revista Latinoamericana de Matemática Educativa,* 13(1), 89-112.

Rogalski, M. (2000). The Teaching Experimented in Lille. In J.-L. Dorier (ed.), On the Teaching of Linear Algebra. Dordrecht : Kluwer, pp. 133-149.

Sierpinska, A. (1994). Understanding in Mathematics. Routledge: The Falmer Press.

Trigueros, M. & Ursini, S. (2003). Starting college students' difficulties in working with different uses of variable. In A. Selden, E. Dubinsky, G. Harel and F. Hitt (Eds.) Research in Collegiate Mathematics Education V, *CBMS Issues in Mathematics Education*, v. 12. American Mathematical Society/Mathematical Association of America, pp.1-29.

Trigueros, M. & Oktaç, A. (2005a). La Théorie APOS et l'Enseignement de l'Algèbre Linéaire. Annales de Didactique et de Sciences Cognitives, 10, 157-176.

Trigueros, M. (2005b). La noción del esquema en la investigación en matemática educativa a nivel superior. Educación Matemática 17 (1), 5-31.

Trigueros, M., Oktaç, A. & Manzanero, L. (2007). Understanding of Systems of Equations in Linear Algebra. Demetra Pitta – Pantazi & George Philippou (Eds.), Proceedings of the 5th Congress of the European Society for Research in Mathematics Education, CERME (pp. 2359-2368).

Student approaches and difficulties in understanding and use of vectors

Oh Hoon Kwon
Department of Mathematics
Michigan State University
kwonoh1@msu.edu

A configuration for analyzing vector representations based on multiple representations, semiotic representation, cognitive development, and mathematical conceptualization, to serve as a new unifying framework for studying undergraduate student approaches and difficulties in understanding and use of vectors is proposed. Using this configuration, the study will explore five important transitions: physics to mathematics, arithmetic to algebraic, analytic to synthetic, geometric to symbolic, concrete to abstract, and corresponding student difficulties along epistemological and ontological axes. As a part of validation of the framework, a mini-study on undergraduate students' approaches and difficulties in understanding and use of vectors is introduced, and we see how useful this new framework is to describe and analyze student approaches and difficulties in understanding and use of vectors.

Key words: representation, vector, vector representation

Vectors are very useful tool for solving real world questions mathematically. Vectors are applied widely to various fields in natural science, engineering and mathematics, even in social science and economics. They also have a valuable role in mathematics itself regardless of any relationship to real world and still have its own significance in advanced mathematics. Undergraduate students usually experienced vectors in school physics and school mathematics. When students study undergraduate mathematics, they see vectors again in multivariable calculus, linear algebra, abstract algebra and geometry courses. Some students see vectors in introductory physics or engineering courses while they are studying vectors in mathematics. Although undergraduate students' experiences with the concept of a vector varied, students still have difficulties in understanding and use of vectors.

In this research, we explore the following: (1) constructing a framework to analyze student approaches and difficulties in understanding and using of vectors, (2) classifying approaches and difficulties, (3) seeing how much one approach prevails over the others in student thinking, and (4) looking for the partial sources of student difficulties.

Root of the Framework for Vector Representations

Most of the studies about multiple representations are centered at the concept of a function (Janvier, 1987; NCTM, 2000). In a case of functions, there could be verbal, table, graphical, symbolic representations and they are usually not hierarchical. Moreover, one can solve one problem with various representations. For example, finding x-intercepts of a function can be done with a table, a graph, a symbolic form, even with verbal representation. Having a question of functions, the use of multiple representations is worth to try because sometimes a different representation gives a different and new perspective to look at the problem with and to look for the answer that can help students' further understanding. Unlike the representations of a function, vector representations have a hierarchy and are strongly dependent on the contexts of given questions. To grasp what student approaches and difficulties are in understanding and use of vector representations, many different contexts and different levels of sophistication should be considered.

Vector representations have a strong hierarchy so that one cannot fully enjoy the privileges and the benefits of multiple representations of a vector until understanding enough levels of sophistication. Representations of a vector are strongly dependent on the contexts of given questions. If a question asks a proof of synthetic geometry, only a few representations can be applied.

Many mathematics teachers and professors already knew student difficulties from their experience of teaching. However, those difficulties are not classified systematically and they are very scattered and isolated. As Tall (1992) mentioned, “the idea of looking for difficulties, then teaching to reduce or avoid them, is a somewhat negative metaphor for education. It is a physician metaphor - look for the illness and try to cure it. Far better is a positive attitude developing a theory of cognitive development aimed at an improved form of learning.” To have more positive attitude, we need to have deeper understanding of student approaches and difficulties on vector representations to the level of the theory of cognitive development.

This necessity of reflecting cognitive development in representations arises in two different directions of study. According to Duval (2006), a simple definition of representation such as “something that stands for something else” can be interpreted from two different perspectives: mental representation theory centered at internal, external communications and mathematical knowledge acquisition; semiotic representation theory that focus on signs and their associations produced according to rules. However, Duval (2006) tried to combine them together and studied semiotic representations at the level of mind’s structure. His identified sources of incomprehension from semiotic systems of representation that have three different purposes: to designate mathematical object, to communicate, and to work on/with mathematical object. He insisted that a transformation from a sign to the other, or a substitution of some semiotic representations for another should be considered with three components of semiotic representation system: representation content, semiotic register used, and represented object. The difference between physics representation and mathematics representation of vectors is gleaned from these components.

Most studies about vectors are from physics point of views related with physical quantities and by physics educators. J. Aguirre and Erickson (1984); J. M. Aguirre (1988); Knight (1995); Nguyen and Meltzer (2003) are just a few of them. Some studies such as Watson and Tall (2002); Watson, Spyrou, and Tall (2003) attempted to analyze student approaches and difficulties on vector representations with more mathematical point of views. However, their studies cover only secondary level mathematics and the transition from physical thinking to mathematical thinking. This brings up a necessity of the new framework for investigating vector concepts that can cover vectors in more advanced and wider ranges of undergraduate mathematics as well as in physics and secondary level mathematics.

Student approaches and difficulties in learning and using of vectors in undergraduate mathematics are very complex issues that have not yet definitely resolved. Dorier (2002) brought up these issues and analyzed them with a series of research. However this book placed the focus at linear algebra so that vectors in geometry or other subject fields in mathematics were covered very briefly. Linear algebra is just one of the fields that requires the concept of vectors frequently, but most studies on the concept of a vector so far are regarded as parts of linear algebra (Dorier, 2002; Harel, 1989; Dorier & Sierpinska, 2001).

Lesh, Post, and Behr (1987) pointed out five outer representations including real world object representation, concrete representation, arithmetic symbol representation, spoken-language representation and picture or graphic representation. Among them, the last three are more abstract and at a higher level of representations for mathematical problem solving (Johnson, 1998; Kaput, 1987). However, in most cases, picture representation is not geometric enough to show geometric structures, and graphical representation does not reflect synthetic geometry point of views but rather reflect analytic geometry point of views.

The problem of vector representations lies not only on the multiple representations but also on the translations. Sfard and Thompson (1994); Yerushalmy (1997) are based on the assumption that students ability to understand mathematical concepts depends on their ability to make translations among several modes of representations. Tall, Thomas, Davis, Gray, and Simpson (1999) analyzed several theories of these. These translations or transitions are referred to as *encapsulation* by Dubinsky (1991) and *reification* by Sfard (1991). The work of Sfard (1991) appears to view both operational and structural conceptions as important in mathematical understanding. Sfard focused on an operational/structural duality of mathematical conceptions. A structural conception enables recognition (at a glance) and manipulation as a whole; an operational conception is grounded in actions, processes, and algorithms. Sfard contended that the development of students conceptions can be viewed as occurring in three stages; interiorization, condensation, and reification; and that "to expect that a person would arrive at a structural conception without previous operational understanding seems... unreasonable" (Sfard, 1991) The proposed configuration of vector representation reflects this idea of encapsulation or reification not just in symbolic modes of representation but also in geometric modes of representation that has not been studied much along with algebra view point (Meissner, 2001b, 2001a; Meissner, Tall, et al., 2006). One important progress is that this configuration of vector representations can capture the whole picture of encapsulation or reification both happen in a symbolic way and a geometric way simultaneously.

Construction of a Framework

This research focuses on issues arising when representations for vectors are utilized in undergraduate mathematics instruction. Ultimately, the issues we plan to investigate in the future include the following:

1. What student difficulties in understanding and use of vector representations can be identified in the undergraduate mathematics curricula?
2. What generalizations might be possible regarding the relative degree of difficulty of various representations in learning vector concepts? That is, given a class and content do some forms of commonly used representations generate a large number of difficulties?

To investigate the above issues, we propose a new framework combining multiple representations, semiotic representation, cognitive development and mathematical conceptualization in order to understand vector representations. This framework serves as a new unifying framework for studying student approaches and difficulties in understanding and use of vectors. Using this configuration, the study explores five important transitions: (A) physics to mathematics, (B) arithmetic to algebraic, (C) analytic to synthetic, (D) geometric to symbolic, (E) concrete to abstract, and corresponding student difficulties along epistemological and ontological axes. As Zandieh (2000) stated in her study on the framework for the concept of a function, "The framework is not meant to explain how or why students learn as they do, nor to predict a learning trajectory. Rather the framework is a 'map of the territory,' a tool of a certain grain size that we, as teachers, researchers and curriculum developers, can yield as we organize our thinking about teaching and learning the concept..." this new framework serves as a 'map of the territory'.

Introduction of the Configuration

Figure 1 is the configuration for analysis of student approaches and difficulties in understanding and use of vectors. It has two axes, ontological axis and epistemological axis. Ontology is the philosophical study of the nature of being, existence or reality in general, as well as the basic categories of being and their relations. Epistemology is the branch of philosophy concerned with the nature and scope (limitations) of knowledge. Vector representations can be classified into those two aspects. For example, geometric or symbolic representations of vectors are the matters of existence or being. Concrete vs. abstract characteristic is also related with ontological aspect of vectors. On the contrary, arithmetic to algebraic transition and analytic to synthetic transition are more related with epistemological aspect. Both ontological axis and epistemological axis provide a good way to locate and see a vector as a mathematical object on undergraduate mathematics.

Figure 1. Configuration of vector representations

The first quadrant of the configuration is related with mathematics. The third quadrant is more related with physics. The origin represents two important jumps from physics to mathematics. Ontologically, the origin is a shift of a view, from vectors as representations of physical quantities with physical units to vectors as representations of mathematical objects. Epistemologically, this origin is a shift related with understanding of mathematical equivalence relations. In the domain of mathematics, each axis has two important jumps that can be easily identified. Those jumps will be explored in the next sections. Ten different representations of vectors are identified in this configuration.

Arrows on a grid. This is a representation of a vector in physics. In physics, a vector is a quantity with direction and magnitude. To represent a vector, they usually use an arrow. In this representation, physical directions and physical magnitudes are important. Without Cartesian axes, physical directions, for example NSEW, can be used and physical magnitudes can be corresponding to the lengths of the arrows with some physical units that can be measured by a grid. Therefore measuring the length of the arrow is significant in this representation. Arrows could be on a plane or on a space with axes as well. However, in physics representation, we see objects or points with vectors as things that are not actually on a plane or on a 3D space, but are within a plane or within a space moving within a plane or within a space.

Arrows with axes and scale. This is the first representation of a vector in mathematics, closest to a physics representation. Once an arrow is put on the Cartesian coordinates, it loses its physical attachments such as physical directions or physical magnitudes. Instead, the direction of a vector can be described by the origin and a point away from the origin measured by the scales on each axis. The length of the arrow defines the magnitude. It does not have any physical units. We see the points or object as things on a plane or a 3D space.

Arrows with axes but no scale. This representation is very similar to the representation above. The only difference is that the exact length of an arrow is not important here but the relative length is important. Measuring the length of a vector at this stage is useless. However the role of the origin as a starting point of position vectors is still important. At this stage, even though a geometric object such as a triangle or a polygon lies on the coordinate plane, it will be interpreted as a collection of points.

Arrows without axes or scale. Because there are no axes and the origin at this stage, the direction and the magnitude of a vector are not critical issues. Instead, the structure of how several vectors are related is important. For example, three vectors on a triangle can produce a special structure of sum or difference of vectors. Angles between vectors, parallel, perpendicularity are essential features at this stage.

Numerical column/coordinate form. Once a basis is given, the tip of a mathematical position vector can give a set of numbers that specify the vector. That set of numbers can be described as a column or coordinate form depending on whether it is from the standard basis or not. At this stage, a vector lives in two or three-dimensional real vector space with an inner product.

Column/coordinate form with variables. Instead of a set of numbers, we use two or three variables and functions. For example, $\begin{pmatrix} a \\ b \end{pmatrix}$, (x,y), $(f(x),g(x),h(x))$. This form of representation is very useful for describing parametrized curves or surfaces in multivariable calculus.

Reduced symbols. At this stage, a single letter or two can represent a column vector with numbers or variables as well as an arrow. For example, $\overrightarrow{AB}$, or $\vec{u}$. It makes the structure of operations of vectors easier to be seen, but it is meaningless to calculate actual results with numbers. The length of a vector exists only in a symbolic form, not given in numerical values. However there is still strong connection to geometric representations so that the length is related with the actual geometric length.

Numerical n-tuple. This is just a four or above dimensional extension of numerical column/coordinate form. Because it is hard to connect a vector with an arrow with axes and scale, it can be viewed as abstract form. However, if one gives up the visual orthogonality of basis vectors, one can still possibly connect arrow to a vector in an abstract way. And at this stage, one can actually calculate the length of a vector by simple extension of the way measuring the length of a vector in two or three-dimensional real vector space with simple extension of Pythagorean theorem. This length is more abstract and constructed by the definition of the inner product of a vector space.

N-tuple with variables. Using variable entries, we can easily extend numerical n-tuple to n-tuple with variables.

Abstract elements in a vector space. This representation does not have any information of a single vector. It is used to describe the structure of an abstract vector space. Zero vectors, inverse vectors are defined in an abstract way. Axioms define a vector space.

Development of the Configuration

Observing the way the concept of vectors is described in high school textbooks, college textbooks and various research; observing and listening to the way that mathematics education researchers, mathematicians, graduate students in mathematics and mathematics education, and undergraduate students describe the concept of vectors were how this framework developed.

According to Hillel (2002), undergraduate linear algebra courses generally included three modes of description of the basic objects and operations. These three modes of description co-exist, were sometimes interchangeable, but are not equivalent. They were: the abstract mode using the language and concepts of the general formalized theory, the algebraic mode using the language and concepts of the more specific theory of $\Re^n$, the geometric mode using the language and concept of 2- and 3-space. Using these three description modes, Hillel (2002) classified the representations of a vector. In the abstract mode, a vector is an element in a vector space defined with axioms. In the algebraic mode, a vector was an n-tuple of numbers in $\Re^n$. In the geometric mode, a vector was a directed line segment, or point. Within each mode, vectors, vector operations and transformations had particular depictions, terminology and notation, and there were mechanisms that enable one to move from one mode to another. Hillel (2002) also subdivided the geometric mode into three levels such as coordinate-free geometry level, coordinate geometry level, and vector-as-point level. Sierpinska (2002) also classified students' thinking and reasoning in linear algebra courses into three modes. Synthetic-geometric, analytic-arithmetic, and analytic-structural modes of thinking and reasoning were those. This classification was different from Hillel (2002)'s classification in a way that this was more focused on students' thinking and reasoning than the descriptions from history or epistemology. Even if a student was working on vectors in $\Re^n$ that were the algebraic mode of description, he or she could use many different modes of thinking and reasoning to figure out the problem, the situation, and the solution. Each mode was not independent from the others. They co-existed and sometimes were interwoven with each other.

These classifications of the concept of a vector were very useful to understand what vectors were in mathematics, but not helpful to see the relationship among representations, the translations and what student difficulties were in learning the concept of a vector and use the vectors in various contexts. Pvalopoulou's work (as cited in Dorier, 2002) as an application and verification of Duval's theory of semiotic representation in the context of teaching linear algebra covered this relationships. The author distinguishes between three registers of semiotic representation: the graphical register (arrows), the table register (columns of coordinates), and the symbolic writing register (axiomatic theory of vector spaces). She proposed problems in one register and asked for translations into another imposed register. This brings out the asymmetric direction of the conversion activity: 7% success in converting from the table register toward the symbolic register, 72% for the opposite conversion, 5% success in converting from the graphical register toward the symbolic register, 40% for the opposite conversion, 83% success in converting from the 2D table register toward the graphical register, 34% for the opposite conversion, and 35% success in converting from the 3D table register toward the graphical register, 68% for the opposite conversion. These results gave warrants for the structure of the configuration and the existence of the levels of conceptualization in the configuration despite the graphical register did not cover the full geometric representations on the configuration.

These studies set up their goal as understanding vectors for linear algebra, so that the important part of vector use in geometry was somewhat neglected. Some students entered a first course in undergraduate geometry with a significant amount of previous experience of vectors in mathematics courses such as multivariable calculus or linear algebra where the representations of vectors were often symbolic. Geometric representations were not developed well enough in those courses. This may well serves as a stumbling block to use the geometric representations of vectors as a way to understand geometry. As a result, students may implicitly believe that use of vectors in geometry is less than desirable or vector geometry is just related with analytic (coordinate) geometry not related with synthetic (coordinate-free) geometry. This inclined scaffolding of vector representations should be overcome in both teaching and learning of vector concepts. Therefore, we needed to organize and structure a wide range of possible understanding and use of vector representations covering beyond linear algebra as levels of conceptualization of vectors across all the fields of mathematics.

Features of the Framework

This new framework has some important features. First, it suggests that the interplay between ontological aspect and epistemological aspect is critical in understanding and use of vectors and the key transitions between representations require both ontological and epistemological aspects of understanding simultaneously. Second, it can distinguish and put greater emphasis on difference between analytic geometric representations of vectors and synthetic geometric representations of vectors. It can also distinguish and put greater emphasis on difference between physical representations of vectors and mathematical representations of vectors. Furthermore, it distinguishes, shows, embeds, and connects parallel developments of symbolic representations and geometric representations along with cognitive development theories such as reification, or APOS theory. And finally it systematizes the transitions between various representations of vectors.

Validation of the Framework

For the initial validation of the proposed configuration of vector representations, we used several literatures about multiple representations, semiotic representation, cognitive development theories in mathematics education such as process-object encapsulation, reification, three world of mathematics etc., and mathematical conceptualization. And then using the surveys and follow up interviews, we checked if students can solve problems developed based on each stage and transition jump on the configuration and we figured out the difficulties. The main method here for the validation was both quantitative and qualitative data analysis of the surveys and interviews.

Classifying Approaches and Difficulties

Student difficulties in learning and use of vectors in mathematics can be identified by the configuration of vector representations in the following way.

Approaches

Student approaches can be divided into two big categories. Analytic approach is the approach that students uses column/coordinate forms of vectors frequently and decomposes vectors into detail numbers or variables in order to construct algebraic method of solving problems. Synthetic approach is the approach that focuses more on the geometric structures of vectors. Instead of getting into the detail, students who take this approach would solve questions with bigger pictures and relationships behind holistically.

Difficulties along Epistemological axis

The critical transitions here are transition (A) from physics to mathematics, transition (B) from arithmetic to algebra, and transition (C) from analytic view to synthetic view. The epistemological aspect of transition (A) is related with equivalence relation in the definition of a vector as a directed line segment. Without understanding of mathematical equivalence relation, one cannot understand of the concept of free vectors or mathematical position vectors that are very important for ontological shift from the geometric representation to symbolic representation. Usually this transition happened in high school mathematics. Watson et al. (2003); Poynter (2004) explained the effective way of transition from the embodied world of mathematics to the proceptual world of mathematics that is corresponding to this transition (A) and more transitions like transition (D) on the configuration. But they were not actually able to see what problems were. For example, their resolution or decomposition example is not related with students' understanding of equivalent vector concept from advanced mathematics. Their "action-effect" ways of teaching can mislead the translation by a vector in such a way that translation vectors are very limited to vectors nearby object and its image.

Transition (B) from arithmetic to algebra occurs actually in the earlier mathematics to students. Transition from arithmetic to algebra is one of the main topics in elementary and middle school mathematics. Even this happened earlier, we still can see if college students actually have that variable thinking in this higher level mathematics with different context. In college level, the major transition here will be transition (C) from analytic view to synthetic view, because it has an ontological shift such as reification or process-object encapsulation in it even though it is an epistemological jump. For example, a geometric object can be studied both analytically and synthetically. Switching back and forth of the ways of thinking and investigating a geometric object is epistemological shift. However, when we use vectors to investigate a geometric object, it is easily seen that corresponding transition from analytic representation to synthetic representation of vectors is related with this epistemological shift as well as ‘process object encapsulation’ that is classified as an ontological shift. To see the difficulty for the major transition, one can give problems of vector additions and subtractions. If one has gone through process object encapsulation or reification, especially in geometric representations, he or she can think vector additions and subtractions in a structural way. He or she will not do calculations by drawing parallelograms, inverse vectors or by measuring the magnitudes. Instead, he or she will see the answer directly from the triangles or polygons and direction of vectors on them.

Difficulties along Ontological axis

There are three transitions here, transition (A) from physics to mathematics, transition (D) from geometric object to symbolic object, and transition (E) from concrete object to abstract object.

Transition (A) from the ontological aspect is about the transition from vectors as representation of physical quantities that have both directions and magnitudes to vectors as mathematical objects that can be represented by directed line segments. Once this transition is done, vectors will not have any physical attachments such as physical directions or units.

Transition (D) is important transition on this axis, because in high school level, with standard basis vectors, this transition is not hard to be achieved. The coordinates for the end points of position vectors will automatically be the coordinate form of vector representations or become column vector forms. However, at undergraduate mathematics level, this transition implies understanding of the difference between vector space and Euclidean coordinate space. The coordinate of the points will not directly correspond to the representations but the component-wise observation on linear combination of basis vectors will give correct coordinate representations of vectors on Euclidean space. This can be the major transition at undergraduate level mathematics, because it has an epistemological shift such as understanding of basis vectors in it even though it is an ontological jump.

Transition (E) is somewhat easier than the rest if it is restricted in symbolic extension, because symbolically it is just appending one more component on the representation. However, strong connection between geometric representations and symbolic representations sometimes will not help students generalize the vectors to higher dimensions such as four or above. And to understand abstract concept of vectors, it is needed to understand the epistemological role of the inner product in vector space in this transition. In 2D or 3D, both geometrical and symbolic representations assume that the magnitudes of two vectors determine the inner product of two vectors. In 4D or above, or more abstractly, the inner product determines the magnitude of a vector. It is a big transition from concrete concepts to abstract concepts of vectors.

Other difficulties on the first quadrant

Other than above difficulties, difficulties that occur on the mixed stages of the configuration are more complex. If one assumes that the stages are just reflections of epistemological shifts, ontological shifts and their simple mixes, then ‘column/coordinate form with variables’ will be easily achieved by understanding ‘arrows with axes but no scales’ and ‘numerical column/coordinate form’. However it is really doubtful and hard to see due to complexity.

Mini Study on Student Approaches and Difficulties

Research Questions

By proposing the configuration of vector representations, we hypothesize that students difficulties lies on ontological and epistemological jumps on transitions in the configuration of vector representations, and students tend to use particular representations more and confine their understanding and use of vectors in a few approaches rather than have flexibility in understanding and use of various approaches. Unlike the assumptions from various studies that students have more difficulties in symbolic representation than geometric representation, some students have more difficulties in geometric representations vectors than symbolic representations. Hence, the following will be the research questions that we will investigate in the main study:

(1) What student approaches and difficulties can be identified in understanding and using of vectors?
(2) How useful is this framework to describe undergraduate students' understanding and use of vectors?

As a preliminary report, we explore some blurry snapshots of these in this paper.

Method

This study was conducted during the 2010 Fall Semester at a large, Midwest public university. A task based survey questions that ask student background on vectors, student approaches in representations, and difficulties was given. This portion of the study attempts to focus on two aspects of student performance in solving problems with various vector representations. One aspect will be identifying the representation stages in the configuration that students are able to use by checking five important transitions. The other aspect will be identifying student difficulties located in the transitions among representation stages in the configuration. To explore the configuration, survey data on each transition will be analyzed with descriptive statistics. In order to see the usefulness of the proposed configuration of vector representations, the follow up interview was also conducted. The main purpose of the interview was to check and modify survey questions so that they can capture more effectively student difficulties along with transitions on the configuration. After administering the actual survey, students were selected for interviews. The selected students signed up for an one-hour block of time for their interview on the day and time that was most convenient for them. Interviews were held in a neutral location away from the students' classrooms and were audio-recorded for further analysis. Transcribed interviews were coded and analyzed with the constructed framework and semiotic representation theory.

Participants and Protocol

Twenty-nine students from senior level Capstone in Mathematics course for pre-service secondary mathematics education participated in the experiment. Pre-service teachers are required to study enough advanced level of vector representations. (Engineering students or economics students could limit their understanding of abstract vectors, for example.) 6 questions about their background information such as major, high school courses related with vectors, college courses related with vectors, etc. were asked together with 12 vector questions with topics about transitions, forces, robot arm, origin, non-standard basis, rotations, polygons, a very long sum, cube, triangle midpoint, associativity, point/vector. After the survey, four students were selected for the follow up interviews to probe deeper student thinking.

Results and Discussion

Some interesting results were founded. Among them, this paper will introduce two examples briefly.

First, we could identify *physics representation* in student mathematical thinking of vectors. Students tended to use and think physics representations (pseudo-geometric and pseudo-abstract diagram) even though the questions were nothing to do with physics contexts. Students also had difficulties in transition (A) from physics representation to mathematics representation. When they were asked the following questions about mathematical transformation with translation vectors, they could not connect the equivalence conditions of vectors with vectors representing a geometric translation.

Question 7. Translation: A translation can be represented by a vector $\vec{v}$, $T_{\vec{v}}(P) = P + \vec{v}$ for any point P.

(a) List all vectors that do NOT represent the translation of triangle A to triangle B in the figure.

(b) Circle all vectors that are equivalent to $\vec{a}$.

Figure 2. Physics to Mathematics Transition Question

Student understanding of the relationship between the represented object (geometric translation) and the representation (directed line segment) of it was not strong enough to connect their mathematical knowledge of vector equivalence and translations. They rather saw a directed line segment as a diagram attached to geometric objects that represented directions and magnitudes of the physical motion of the physical objects. In the question 7 (a), $\vec{d}$, $\vec{e}$, $\vec{f}$ are not related with translation. However, some students who were still using physical representation of vectors interpreted that $\vec{g}$ and $\vec{h}$ were not related with this translation as well, because these vectors were away from actual triangles. $\vec{c}$ was also chosen as a vector not related with this translation because its initial point was on the image of the triangle. The following cluster tree diagram from 29 students responses shows this clearer. (See Figure 3.)

In the follow up interview, a student explained what he was thinking when he tried to solve this question 7(a). His answer was $\vec{d}$, , , :

> $\vec{d}$ comes back, it's a translation from B to A, so it's opposite. is a translation of onward, so $\vec{a}$, A going to B would not include , in my mind. is the correct form up 1 over 3 but it's just way up here and there is a, two separate points than A and B and the same with which is the same thing but just sort of from other points… so if you moved it over [nearer to triangles], then yeah it would be…but I guess, in my mind, it's something totally different.

When the interviewer asked about vector $\vec{b}$ with emphasizing that the vector itself did not touch the sides of triangle A, he responded:

> If you shifted this over here to A. $\vec{b}$ is touching the point on it. Very close……

From his response, we can see that this directed line segment (arrow) representation was not understood as a geometric translation that translate all the points on the plane, but rather as a representation of a displacement or physical motion of triangle A. The distance between the geometric object and the directed line segment was the critical condition of decision-making. His idea of translation was described as:

> Translation is just moving it (a triangle) up 1 unit and shifting that triangle over 3 units, Umm, I believe, yeah that's going to be where it's going to go. Yeah so from A to B it would be going up 1 and over 3.

Figure 3. Cluster Tree from Student Responses on Translation Vector

More interesting fact is that most students answered correctly on the second question about vector equivalence. As we can see from the cluster tree diagram below (Figure 4), even though there seems to be some degrees of acceptance as equivalent vectors, $\vec{b}$, , , $\vec{h}$ are clustered as equivalent as $\vec{a}$.

This result tells us that vector representation should be explored by refining the notion of representations and by integrating different theoretical perspectives used to describe cognitive development in mathematics. In particular, this example of student thinking on a vector representation for a geometric translation in Euclidean geometry can be analyzed from the physical embodiment and semiotic representation point of views.

The analysis considers the role of the physical embodiment in the vector representations, and also utilizes the tools of semiotic representations and cognitive development theory with newly developed vector representation framework. For example, if we can refine the notion of register in semiotic representation theory, physics (physical diagram) register as an additional register to well-known registers such as table, graphical, or symbolic registers in vector representations that take into account how students actually think. Actually, the motion of geometric objects in mathematics is quite different from that in physics or realistic context. Freudenthal (1983) described the distinction between physical motion and mathematical motion. Physical motion is something that occurs to an object within space or plane within time, but mathematical motion should be differentiated from physical motion in three ways: from *the limited object* to *the total space (plane)*, from *within* space (plane) to *on* space (plane), from *within time* to *at one blow*. Because this motion of geometric objects cannot be shown fully on paper, students should use representations to describe how an object should be or has been moved. They can invent various ways to represent the motion. Translations move a figure a fixed distance in a given direction. Translation arrows (vectors) represent the distance and direction for moving the figure. However the meaning of the arrows varies from time to time. It is related with student cognitive development and their interpretation of mathematical world (Watson, Spyrou, & Tall, 2003).

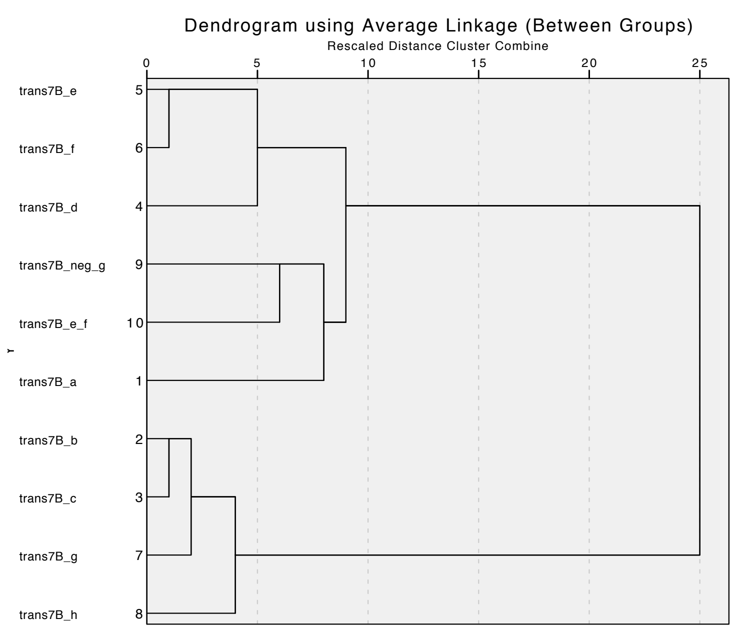


Figure 4. Cluster Tree from Student Responses on Equivalent Vector

The next result we have brought is the *prevalence of analytic approach* in understanding and use of vectors. The following question was related with physics to mathematics, geometric to symbolic, and analytic to synthetic transitions.

Question 16. Triangle Midpoints: In the following △ABC, AB = 2AD and AC = 2AE. We want to show that BC is parallel to DE.

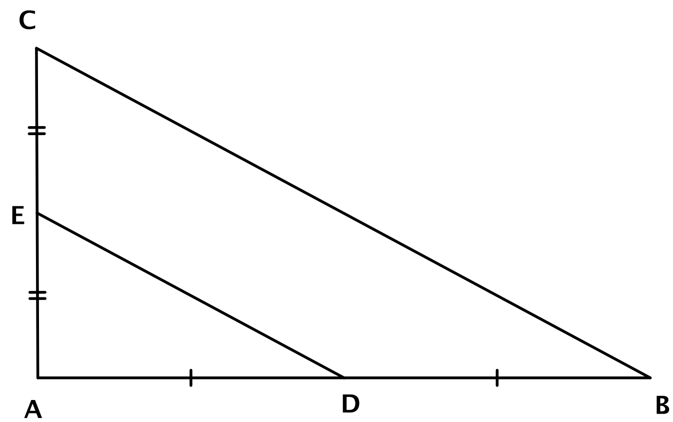


Figure 5. Triangle Midpoints Theorem

Which form of vectors, do you think will be most useful? Circle your answer.

(i) (x, y), (ii) $\begin{pmatrix} a \\ b \end{pmatrix}$, (iii) $\overrightarrow{PQ}$, (iv) $\vec{u}$, (v) Others

(i) and (ii) can be regarded as analytic approach, and (iii), (iv) can be synthetic approach. Because we can put the directed line segment on the sides of triangle and see this question as vector subtraction in the structure sense and scalar multiplication question in the algebraic manipulation sense, synthetic approach is more reasonable. Due to no guarantee on given triangle is right, use of coordinate/column form of vector representation and component-wise calculation through analytic approach is not appropriate in this setting. However, 10 students (34.5%) of 29 students chose (i), and 6 students (21%) chose (ii). 10 students (34.5%) chose (iii), and 0% on (iv). 3 students (10%) gave no responses. Many students attempted this with synthetic approach, but still 55.5% used analytic approach on this question. Their approach was almost about calculating slopes of the sides and tended to put the triangle on a coordinate plane.

As a partial explanation of this, it is believed that U.S. curriculum put more emphasis on the upside down 'L' shape learning path (analytic approach) on the configuration of vector representations whereas European or Asian curriculum put more emphasis on the left handed 'L' shape learning path (synthetic approach) on the configuration of vector representations so that U.S. students would have more difficulties on geometric representations of vectors, even when upside down 'L' shape path on the configuration, can be easily achieved by European or Asian students as well as U.S. students. This requires more comparative research on bigger samples across the countries. From this research, we cannot just say it is because of the curriculum difference without any empirical evidence, and more study about the partial explanation of the difference will be done in the future.

Limitation

The proposed configuration alone as a framework for analysis of student approaches and difficulties only gave us a really blurry idea of what was actually going on in student thinking. It was more helpful to understand the whole range of vector representations that mathematics and mathematics education community required. Holistic and macro view are useful more for learning trajectory or path. In the on-going and future research, we will provide more careful analysis on student thinking using the triple components of semiotic representation, {{representation content, semiotic register used}, represented object} that will focus more on individual thinking and micro view on vector representations. Macro view together with micro view will hopefully coordinate and systemize our better understanding of student approaches and difficulties on vector representations.

References

Aguirre, J., & Erickson, G. (1984). Students' conceptions about the vector characteristics of three physics concepts. *Journal of Research in Science Teaching, 21(5),* 439–457.

Aguirre, J. M. (1988, April). Student preconceptions about vector kinematics. *The Physics Teacher, 26(4),* 212–216.

Dorier, J. (2002). *On the teaching of linear algebra*.

Dorier, J., & Sierpinska, A. (2001). Research into the teaching and learning of linear algebra. *The Teaching and Learning of Mathematics at University Level,* 255–273.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. *Advanced mathematical thinking,* 95–126.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, *61,* 103-131.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures.* Springer.

Harel, G. (1989). Learning and teaching linear algebra: Difficulties and an alternative approach to visualizing concepts and processes. *Focus on Learning Problems in Mathematics*.

Janvier, C. (1987). *Problems of representation in the teaching and learning of mathematics*. L. Erlbaum Associates.

Johnson, S. (1998). What's in a representation, why do we care, and what does it mean? examining evidence from psychology. *Automation in Construction, 8(1),* 15–24.

Kaput, J. J. (1987). Representation systems and mathematics. *Problems of representation in the teaching and learning of mathematics,* 19–26.

Knight, R. D. (1995). The vector knowledge of beginning physics students. *Physics Teacher, 33,* 74–74.

Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. *Problems of representation in the teaching and learning of mathematics,* 33–40.

Meissner, H. (2001a). Encapsulation of a process in geometry. *PME CONFERENCE, 3,* 3–359.

Meissner, H. (2001b). Procepts in Geometry. *European Research in Mathematics Education II, 58*.

Meissner, H., Tall, D., et al. (2006). Proceptual Thinking in Geometry. *RETIREMENT AS PROCESS AND CONCEPT, 165*.

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

Nguyen, N., & Meltzer, D. E. (2003, June). Initial understanding of vector concepts among students in introductory physics courses. *American Journal of Physics, 71(6),* 630– 638. Available from http://link.aip.org/link/?AJP/71/630/1

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational studies in mathematics, 22(1),* 1–36.

Sfard, A., & Thompson, P. (1994). Problems of reification: Representations and mathematical objects. *In Proceedings of the xvi annual meeting of the international group for the psychology of mathematics education–north America,* 1–32.

Tall, D. (1992). Students' difficulties in calculus. *Plenary presentation in Working Group 3,* ICME, 1–8.

Tall, D., Thomas, M., Davis, G., Gray, E., & Simpson, A. (1999). What is the object of the encapsulation of a process? *The Journal of Mathematical Behavior, 18(2),* 223–241.

Watson, A., Spyrou, P., & Tall, D. O. (2003). The relationship between physical embodiment and mathematical symbolism: The concept of vector. *The Mediterranean Journal of Mathematics Education, 1(2),* 73–97.

Watson, A., & Tall, D. (2002). Embodied action, effect and symbol in mathematical growth. *PME Conference, 4,* 4–369.

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

Zandieh, M. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. *Research in collegiate mathematics education IV, 8,* 103–126.

IMPROVING THE QUALITY OF PROOFS FOR PEDAGOGICAL PURPOSES: A QUANTITATIVE STUDY

Yvonne Lai
University of Michigan
yvonnexlai@gmail.com

Juan Pablo Mejía-Ramos
Rutgers University
pablo.mejia@gse.rutgers.edu

Keith Weber
Rutgers University
keith.weber@gse.rutgers.edu

At last year's conference, we presented a qualitative study providing insight into what mathematicians believe makes a good proof for pedagogical purposes based on eight mathematicians' revisions of two proofs (see Lai & Weber, 2010). In this paper, we empirically test four hypotheses generated from last year's study. This year's study provides quantitative support for the claims that mathematicians believe (1) adding an introductory sentence stating the goals of the proof improves its pedagogical quality, (2) formatting key equations in a proof to emphasize their importance improves their pedagogical quality, and (3) unnecessary statements in a proof lowers its pedagogical quality.

Key words: Proof; Proof presentation; Undergraduate mathematics instruction

1. Introduction

1.1. Proof as explanation in collegiate mathematics education

Providing instructional explanations is a fundamental activity in mathematics instruction (Charalambous, Hill, & Ball, in press); explanation is not only a primary means by which teachers convey mathematical subject matter to their students (e.g., Leinhardt et al, 1991), but also a means of establishing classroom norms, illustrating productive metacognitive processes, and representing the discipline of mathematics (Larreamendy-Joerns & Muñoz, 2010; Schoenfeld, 2010). For these reasons, Charalambous, Hill, and Ball (in press) contend that "providing instructional explanations lies at the heart of teaching, for it requires transforming the content in mathematically legitimate *and* pedagogically appropriate ways" (authors' emphasis). There has been a great deal of research on instructional explanations, largely focusing on teachers' difficulties or inability to provide adequate explanations (e.g., Ball, 1988; Inoue, 2009; Leinhardt, 1989; Lo et al, 2004; Thompson & Thompson, 1994, 1996; Thanheiser, 2009), and more recently on improving teachers' abilities to provide high quality explanations (e.g., Charalambous, Hill, & Ball, in press; Kinach, 2002; Inoue, 2009; Thanheiser, 2010). However, most of this work has been done with elementary mathematics teachers and little work of this type has been done with teachers of tertiary mathematics.

In this paper, we focus on teachers' pedagogical practice in advanced mathematics courses at the tertiary level. In these courses, the predominant way of presenting mathematical subject matter to students is via mathematical proof. By mathematical proof, we mean "a formal and logical line of reasoning that begins with a set of axioms and moves through logical steps to a conclusion" (Griffiths, 2000, p. 2). We note further that other characteristics of this genre include the use of precise definitions rather than informal descriptions of concepts, the (relative) lack of diagrams and other intuitive representations of concepts, and the use of logical syntax (e.g., Weber & Alcock, 2009).

Although many mathematics educators and mathematicians question whether these types of proofs are an appropriate way of conveying mathematics to students (e.g., Davis & Hersh, 1981; Hersh, 1993; Kline, 197; Thurston, 1994), there is little doubt that proof is the predominant way that mathematics is presented to students in advanced mathematics classrooms. For instance, Davis and Hersh (1981) asserted that "a typical lecture in advanced mathematics … consists entirely of definition, theorem, proof, definition, theorem, proof, in solemn and unrelieved concatenation" (p. 151). Dreyfus (1991) claimed that although mathematics instructors may be aware that new mathematics is created through non-rigorous processes, this "does not usually prevent him or her from almost exclusively teaching the one very convenient and important aspect which has been described above, namely the polished formalism, which so often follows the sequence theorem-proof-application" (p. 27). Weber (2004) presented a detailed case study of one professor's teaching that was largely based on presenting proofs of theorems. Based on her observations of three mathematics professors, Mills (2011) found that these professors spent, on average, about half of their lecture time presenting proofs to students.

1.2. What makes a good pedagogical proof? Results from an exploratory study

In a previous paper, we investigated mathematicians' beliefs about what makes a good proof for pedagogical purposes by examining the ways in which eight mathematicians revised two proofs that were to be presented to undergraduate students (Lai & Weber, 2010). Our rationale for conducting this study is that the mathematicians presumably believed their revisions improved the proof for pedagogical purposes, either by removing an aspect of the proof that was undesirable or introducing a desirable feature into the proof. By analyzing what types of revisions that mathematicians made, as well as attending to their justifications for making these revisions, we could gain insight into what characteristics mathematicians believed made a good proof for pedagogical purposes and what mathematical and pedagogical reasons they had for these beliefs. Hence, through the analysis of our data, we believed we could form grounded hypotheses about what mathematicians valued in a proof and why. Four of the hypotheses that we generated our presented below:

- (H1) Mathematicians believe that a proof may be more easily understood by undergraduates when it contains hypothesis and conclusion statements that make explicit what is being accomplished in the proof.
- (H2) Mathematicians believe that emphasizing the main ideas used in a proof can improve its quality for pedagogical purposes.
- (H3) Mathematicians believe that adding extra justifications to support an assertion can improve the clarity of a proof if that justification might be difficult for a student to infer on their own.
- (H4) Mathematicians agree that including unnecessary, irrelevant computations or assumptions in a proof will lower its pedagogical quality.

We generated (H1) because all eight mathematicians in Lai and Weber (2010) added an introductory or concluding statements to at least one of the proofs. We generated (H2) because several mathematicians in Lai and Weber (2010) introduced formulas into the proofs they read because they felt these formulas represented the "key idea" of the proof or were "the heart of the matter". Further, four participants centered what they believed to be pivotal formulas in the proof to emphasize their importance. We generated (H3) because participants frequently added justifications to the proof if they felt students would have difficulty generating these justifications for themselves. We generated (H4) because one of the most common revision mathematicians in Lai and Weber (2010) made was removing statements that were not germane to the main ideas of the proof or were inferences that they believed would be trivial for undergraduates to make.

1.3. Research questions addressed this study

We believed the findings reported in Lai and Weber (2010) were interesting and potentially valuable. However, due to the qualitative and exploratory nature of our study, as well as the small sample size that we employed, we were obligated to be cautious about the generality of our findings. This is why we refer to the findings from Lai and Weber (2010) as hypotheses.

We believe that small-scale qualitative studies are indispensable in mathematics education, both for the development of theory and for the generation of hypotheses. However, we also contend that these types of studies should be the starting point of an investigation, not the ending point. In particular, we believe the generation of useful hypotheses is a crucial function of small-scale qualitative studies, but these hypotheses need to be rigorously tested to ensure their validity and generalizability. In this paper, we test hypotheses (H1), (H2), (H3), and (H4) in a large-scale quantitative study.

2. Related literature

2.1. The communicative functions of proof

A significant function of proof in mathematics is to provide conviction that an assertion is true (e.g., Harel & Sowder, 1998). However, in a seminal paper, de Villiers (1990) argued that proof is much more than that. To mathematicians, proof serves many functions beyond that of providing conviction. Proof can also be used as a tool to explain why a theorem is true, systematize a mathematical theory, or discover new theorems. de Villiers (1990) further argued that an important function of proof is communication. Providing mathematicians with a shared language and standards of argumentation facilitates debate about sophisticated mathematical ideas. Many educators have argued that proof should provide similar roles in the mathematics classroom; proof should not only be used to convince students that a theorem is true, but also to provide explanation and facilitate communication (e.g., Alibert & Thomas, 1991; Hanna, 1990; Healy & Hoyles, 2000; Knuth, 2002). These researchers also lament that proof often is not used for explanation and communication in these classrooms, consequently playing only a limited role in mathematics teaching.

2.2. Educational research on proof presentation

Although there have been few systematic studies on undergraduates' comprehension of proof, both mathematicians and mathematics educators have remarked that students generally find proofs confusing and learn little from reading them (e.g., Alcock, 2010b; Davis & Hersh, 1981; Hersh, 1993; Leron & Dubinsky, 1995; Porteous, 1986; Rowland, 2001; Thurston, 1994). Although there are many reasons why students might learn little from reading proofs, many argue that it is the linear and formal nature of proof that inhibits' students' understanding. Specifically, proofs often mask the intuitive representations that are needed to generate and understand them, and the jargon, logical syntax, and abstractness of a proof are intimidating barriers to comprehension (e.g., Hersh, 1993; Kline, 1977; Leron, 1983; Rowland, 2001; Thurston, 1994). Consequently mathematics educators propose different ways to present mathematical information to students other than formal proof (e.g., Alcock, 2009; Hersh, 1993;

Leron, 1983; Rowland, 2001), although we are not aware of any empirical studies that demonstrate the efficacy of these instructional suggestions. Aside from these instructional recommendations, research on proof presentation in mathematics education has been limited. In particular, we are not aware of any studies on what mathematicians believe are good proofs for pedagogical purposes or how mathematicians choose to present proofs to their students. The studies presented in this paper address these issues.

The primary aim of this paper is to investigate what mathematicians—the ones who usually teach advanced mathematics courses—think constitutes a good mathematical proof for pedagogical purposes. This fills two underrepresented areas of research in mathematics education. First, as Speer, Smith, and Hovarth (2010) argued, there is little research in mathematics education on how mathematicians actually teach. Second, based on a systematic review of the mathematics education literature on argumentation and proof, Mejia-Ramos and Inglis (2009) found that empirical research on proof has mainly focused on students' construction and evaluation of proof; their sample did not contain any empirical studies on how mathematical instructors (or for that matter, students) chose to present mathematical proofs. There are two reasons for undertaking this line of research. First, understanding mathematicians' beliefs about what constitutes a good proof for pedagogical purposes provides researchers with a better lens to interpret mathematicians' pedagogical behaviors, including their in-class behavior as well as how they prepare lecture notes for class. Second, as Alcock (2010b) noted, if mathematicians and mathematics educators are going to engage in a conversation about how to improve collegiate mathematics education, then it is necessary for mathematics educators to understand mathematicians' beliefs and values and to take these beliefs and values seriously. Mathematicians most likely will not adopt innovative instruction that is at variance with their beliefs and values.

3. Methods

3.1. Participants

We recruited mathematicians to participate in this study as follows. Thirty secretaries from mathematics departments in the United States were contacted and asked to distribute an e-mail to the mathematics faculty, post-docs, and Ph.D students of that department. These mathematics departments had strong national reputations within the United States. The e-mails that the secretaries sent to the recipients explained the study they were going to complete and invited them to visit a website with the study if they were interested. A total of 110 participants agreed to participate.

3.2. Validity of internet-based studies

We adopted an internet-based study to maximize our sample size using a methodology similar to Inglis and Mejia-Ramos (2009). The validity of internet-based experiments was studied by Krantz and Dalal (2000), who compared 20 internet-based studies with their laboratory equivalents and found a "remarkable degree of congruence" between the methodologies. A similar conclusion was reached by Gosling et al. (2004). Given these findings, and the impracticality of obtaining large samples of research-active mathematicians, we believe our methodology is justified.

3.3. Materials.

The materials used in this study are presented in the Appendix. Participants were shown a Master Proof . They were asked to compare this Master Proof to five other proofs (M1, M2, M3, M4, and M5) that had modifications to the Master Proof in blue. The participants were asked to judge whether these modifications increased or decreased the pedagogical value of the proof.

The first proof, M1, was used to test our first hypothesis, H1. It added an introductory and concluding sentence to the Master Proof to make explicit the proof framework being employed. M2 was used to test H2 by reformatting two important formulas in the proof to highlight their significance. M3 was used to test H3 by adding a justification in the proof that we felt might be difficult for students to infer. If H1, H2, and H3 were correct, we would expect mathematicians to view M1, M2, and M3 respectively as improvements to the proof. M4 and M5 were used to test H4. M4 added an irrelevant calculation to the proof, while M5 added an unnecessary assumption before applying the Mean Value Theorem. If H4 is correct, we would expect mathematicians to view M4 and M5 as lowering the pedagogical quality of the proof. We included two items, M4 and M5, to test the possibility that participants would view any additions or changes to the text positively.

3.4. Procedure

When participants visited the experiment website, they were first asked to indicate whether they were a Ph.D student, a post-doc, or mathematics faculty, as well as their level of experience. They were then shown the Master Proof and told that they would be shown modifications to the Master Proof; they would then be asked to judge whether the changes made the proof “more or less understandable to a second- or third-year undergraduate student”. Next, the participant was presented with a screen containing the Master Proof at the top of the screen and a modified proof beneath that. They were asked if the modified proof made the proof “significantly better”, “somewhat better”, “the proofs were the same”, “somewhat worse”, or “significantly worse”. We coded these responses as 2, 1, 0, -1, and -2 respectively. Along with their evaluations, the participants were given the option of commenting on their responses. For each modified proof, we performed an open coding of the participants’ responses in the same manner in which we coded participants revisions in the revision task. This process was repeated until the participants evaluated all five modified proofs. The order in which the modified proofs were presented was randomized by participant.

4. Results

A repeated measures ANOVA revealed a main effect on participants’ evaluations by which proof the participants evaluated ($F(4, 327) = 231.7$, $p<0.001$), indicating participants did not all believe the modifications were of equal quality. Of the 110 participants, 20 were mathematics faculty, 13 were post-docs, and 77 were Ph.D students. ANOVAs comparing the status and experience of the participants with their response patterns did not yield significant effects ($F < 1$, $p > 0.5$) in both cases, indicating that there was not a significant difference between how mathematics faculty members, post-docs, and Ph.D students performed on this task. Table 1 summarizes the main quantitative results from this study.

Table 1. Summary of data

Condition	Mean score	# participants who thought proof was better	# participants who thought proof was worse
M1	1.29*	97	4
M2	1.05*	88	2
M3	0.02	41	40
M4	-1.66*	6	98
M5	-1.12*	7	94

* Indicates a mean score statistically different than zero with $p<0.001$.

(Note participants who claimed the original and modified proofs had the same pedagogical value were not included as either participants who thought the proof was better or who thought the proof was worse).

4.1. M1: adding an introductory and concluding sentence

The results in Table 1 confirm H1—participants overwhelmingly believed that adding an introductory sentence that makes the framework of a proof explicit improves the pedagogical clarity of the proof. A summary of participants' comments is presented in Table 2. Of the 36 participants who evaluated the proof positively and volunteered to leave comments, 30 cited the benefit of giving students a "roadmap" to the proof and letting the students know explicitly what was being proved. Eight participants argued that the first sentence was good because it reminded the participants what the definition of injectivity was. These comments also reveal that some participants did not view the concluding sentence as an improvement, suggesting the main benefit of the changes lied in adding the introductory sentence.

Table 2. Comments from participants who evaluated M1 positively

Type of Response	Number of Responses (36 total)	Representative Responses
It is beneficial to state the objectives of the proof	30	"Stating what is to be proven is almost always helpful to students", "I think any proof benefits from a summary of what needs to be shown", "Roadmaps. Great."
The second sentence was less helpful than the first	5	"The last blue part is too much, I think", "The last sentence is superfluous"
It is useful to remind students of the definition of injectivity	8	"Yes! Repeat definitions like mad! Drill them into their heads!", "It's always nice to repeat the definition"

(*Note:* Some responses were assigned to more than one category).

4.2. M2: reformatting two important formulas

The results in Table 1 confirm H2—most participants believed that re-formatting the proof by placing significant equations in the proof on their own line and centered improved the pedagogical quality of the proof. Participants’ comments confirm this finding. However, these comments also suggest another reason for why M2 was judged to be an improvement to the Master Proof: the additional spacing simply made the proof easier to read.

A summary of participants’ comments is presented in Table 3. Of the 31 participants who evaluated the proof positively and volunteered to leave comments, 21 cited the benefits of making the proof less cluttered and easier to read, while ten cited the benefits of emphasizing the important equations of the proof. Whether some mathematicians believe that formatting *any* equation to increase spacing would increase the pedagogical value of a proof, or if this should only be done to important equations, is a question that can be addressed in future research.

4.3. M3: adding a justification in the proof

The results in Table 1 failed to support H3. As Table 4 illustrates, many participants (41) participants believed adding the extra justification improved the Master Proof. However, nearly an equal number of participants (40) felt the addition of the extra justification made the proof worse.

Table 3. Comments from participants who evaluated M2 positively

Type of Response	Number of Responses (31 total)	Representative Responses
The added spacing makes the proof less cramped and easier to read	21	“formulae on their own lines makes reading proofs easier”, “Easier to read”, “The original seemed cramped”, “This is easier on the eyes, making the proof psychologically more appealing, I think”
The formatting draws the reader’s attention to the important ideas of the proof.	10	“I think this makes the proof much clearer. Giving these equations their own line emphasizes that this is where the main points in the proof are happening”, “Displayed equations draw the eye to important claims
Other	5	“Now the proof, once understood, can be recalled at a glance”

(*Note:* Some responses were assigned to more than one category).

Table 4. Comments from participants who evaluated M3 positively or negatively

Type of Response	Number of Responses (31 total)	Representative Responses
Favorable evaluation comments (N=10)		
Including the extra calculation may prevent potential confusion about the last step	8	“I can imagine a student being confused by the last step and this change would make it clearer”, “The new statement is redundant, but it saves time in thinking about why it works”
Other	2	“The extra explanation helps a bit, but it makes the proof Bulkier”
Unfavorable evaluation comments (N=15)		
There is benefit to having students make this inference on their own	2	“I decided it’s better for students to figure out the last step on their own”, “We should push students to complete this much reasoning on their own”
The audience is capable of making this calculation so it unnecessarily lengthens the proof	7	“Maybe ‘worse’ sounds harsh, but it adds more words and if you have proved the Mean Value Theorem, then a deduction like this should be straightforward”, “Clutter by repetition”
The added justification is not written well	6	“I think you should say f(x2) – f(1) is greater than zero, therefore f(x1) <> f(x2)”, “Not a good way to conclude the proof”
Other	2	“I suspect it takes more to read the inserted line than to realize what is being pointed out. It would take a study of actual undergrads to make sure”.

(*Note:* Some responses were assigned to more than one category).

A summary of the comments that participants left for both the positive and negative evaluations is presented in Table 4. Eight participants commented that they approved of the modification, believing the extra justification might prevent confusion. However seven who disapproved of the modification believed that the justification unnecessarily lengthened the proof; they argued that undergraduates should be able to make this inference on their own. This suggests that there may not be agreement amongst mathematicians as to what types of inferences undergraduates are capable of making. Further, these comments suggest a tension that mathematicians may feel when adding a justification; while adding a justification might make a proof clearer to some, it might also make the proof unnecessarily longer to others.

4.4. M4: adding an irrelevant calculation to the proof

The results in Table 1 confirm H4—participants overwhelmingly believed the inclusion of the extra calculation made the proof worse. A summary of the 40 comments left by participants who evaluated M4 negatively are presented in Table 5. Most of these participants noted that the included change made the proof longer unnecessarily, while 15 participants commented that the new equation was likely to be distracting or confusing for the reader.

Table 5. Comments from participants who evaluated M4 negatively

Type of Response	Number of Responses (40 total)	Representative Responses
The extra calculations unnecessarily lengthen the proof	31	"The additional formula makes it longer but not clearer", "That isn't the manipulation used at the end, why do it?", "The addition seems irrelevant"
Including the extra calculation will distract or confuse the student	15	"The blue material just adds another equation for the poor reader to parse", "Students may lose time and energy wondering why this line is there and/or develop misunderstandings"
The added expression divides by f'(3) before justifying that f'(3)≠0	4	"This comes before f'(x3) > 0 in the argument so you shouldn't divide through without indicating reasoning",

Other	3	"This takes the argument on a weird path"

(*Note:* Some responses were assigned to more than one category).

4.5. M5: adding an unnecessary assumption

Finally, the results of this study (presented in Table 1) also confirm H4—most participants viewed adding the assumption that f was a real-valued function in the proof diminished its pedagogical quality. A summary of the 40 comments left by participants who evaluated M4 negatively are presented in Table 6. Twenty of these comments note that adding this extra assumption could distract or confuse the students because it might lead them to try and make sense of why it was included.

Table 6. Comments from participants who evaluated M5 negatively

Type of Response	Number of Responses (40 total)	Representative Responses
The assumption is superfluous for this part of the proof	16	"The inserted detail is not pertinent to that area of the proof so I don't see any benefit to pointing it out", "This is unnecessary", "That's not the issue here, honestly"
Including this assumption is distracting and potentially confusing.	20	"The f was real-valued distracts from the main point", "This statement is distracting to the reader", "The blue phrase diverts attention away from the salient hypotheses"
It was implicitly understood throughout the proof that f was real-valued	21	"We are presumably referring to f being real-valued to appeal to properties of an ordered field. But in this context, it's quite obvious so the change just adds words", "The students most likely assume all functions are real-valued"

(*Note:* Some responses were assigned to more than one category).

5. Discussion

The results of this quantitative study revealed three things. First, the results confirmed H1, H2, and H4, supporting our claims that mathematicians believe proofs for undergraduates should make the proof frameworks explicit, format important equations to highlight their importance, and avoid adding unnecessary calculations and assumptions. Second, the comments left by participants deepened our understanding of why mathematicians valued these things. While adding introductory sentences to a proof can remind students of the meaning of definitions and provide a roadmap for how the proof will proceed, some participants found the concluding sentence to be unnecessary. Although some participants commented that the formatting of equations in M2 highlighted their importance and illustrated the main ideas of the proof, even more participants preferred the formatting because it made the proof less condensed and easier to read, suggesting that spacing, and more generally, visual appearance are important aspects that mathematicians value in proofs for pedagogical purposes. The comments on M4 and M5 illustrate one reason that mathematicians prefer that proofs do not have unnecessary calculations or assumptions. They believe these may distract the reader from the main ideas of the proof and confuse the reader by encouraging him or her to try to understand why they were included.

Finally, the results illustrated that H3—that mathematicians value adding an extra justification to a proof—was more nuanced than we believed. In a sense, the results of our quantitative study are consistent with the qualitative findings we found in Lai and Weber (2010).

In Lai and Weber (2010), several participants suggested adding justifications to a proof to bridge logical gaps that students might find challenging. In the current study, a sizeable minority of participants (41 out of 110 participants, or 37%) viewed the added justification in M3 to improve the pedagogical quality of the proof. However, to our surprise, a nearly equal number of participants believed adding this justification diminished the pedagogical quality of the proof, with some commenting that this inference would be obvious to the audience reading the proof. This suggests that mathematicians may strongly disagree on what level of detail is optimal in the proofs that they present to students, in part because they do not agree on what would be obvious to a student.

An interesting question is whether these mathematicians' beliefs are accurate. That is, do proofs that have the desirable features have concrete pedagogical benefits? For instance, will students who read such proofs understand them better or enjoy them more? These questions will be addressed in future research studies.

References

Alcock, L. (2009). e-Proofs: Students experience of online resources to aid understanding of mathematical proofs. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education*. Available for download at: http://sigmaa.maa.org/rume/crume2009/proceedings.html. Last downloaded April 10, 2010.

Alcock, L. (2010a). Interactions between teaching and research: Developing pedagogical content knowledge for real analysis. In R. Leikin & R. Zazkis (Eds.) *Learning through teaching mathematics*. Springer: Dordrecht.

Alcock, L. (2010b). Mathematicians perspectives on the teaching and learning of proof. *Research in Collegiate Mathematics Education*, 7. 73-100.

Alibert, D. and Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.) *Advanced Mathematical Thinking*. Kluwer: The Netherlands.

Ball, D. L. (1988). *Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education*. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Charalambous, C. Hill, H., & Ball, D. (in press). Prospective teachers learning to provide instructional explanations: how does it look and what might it take? To appear in *Journal of Mathematics Teacher Education*.

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

de Villiers, M. D. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, 17-24.

Dreyfus, T. (1991). Advanced mathematical thinking processes. In D. Tall (Ed.) *Advanced Mathematical Thinking* (pp. 25-41). Dordrecht: Kluwer.

Gosling, S.D., Vazire, S., Srivastava, S., & John, O.P. (2004). Should we trust web-based studies? A comparative analysis of six studies about internet questionnaires. *American Psychologist*, 59, 93-104.

Griffiths, P.A. (2000). Mathematics at the turn of the millennium. *American Mathematical Monthly*, *107*, 1-14.

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

Harel, G. & Sowder, L. (1998). Students proof schemes. In A. Schoenfeld, J. Kaput, E. Dubinsky (Eds.), *Research in collegiate mathematics education III*, (pp. 234-282). Washington, DC: American Mathematical Society.

Healy, L. & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 31, 396-428.
Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24(4), 389-399.
Inglis, M. & Mejia-Ramos, J.P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27, 25-50.
Inoue, N. (2009). Rehearsing to teach: Content-specific deconstruction of instructional explanations in pre-service teacher training. *Journal of Education for Teaching, 35*(1), 47–60.
Kinach, B. M. (2002). Understanding and learning-to-explain by representing mathematics: Epistemological dilemmas facing teacher educators in the secondary mathematics "methods" course. *Journal of Mathematics Teacher Education, 5*, 153–186.
Kline, M. (1977). *Why the professor can't teach*. New York: St. Martin's Press.
Knuth, E. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, 33(5), 379-405.
Krantz, J.H. & Dalal, R. (2000). Validity of web-based of psychological research. In M.H. Birnbaum (Ed.) *Psychological experiments on the internet*. San Diego: Academic Press.
Lai, Y. & Weber, K. (2010). What makes a good proof: Investigating mathematicians' proof revisions. In *Proceedings of the 13th Conference for Research in Undergraduate Mathematics Education.* Available for download at http://sigmaa.maa.org/rume/crume2010/Abstracts2010.htm
Larreamendy-Joerns, J., & Muñoz, T. (2010). Learning, identity, and instructional explanations. In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 23–40). New York: Springer.
Leinhardt, G. (1989). Math lessons: A contrast of novice and expert competence. *Journal for Research in Mathematics Education, 20*, 52–75.
Leinhardt, G., Putnam, R. T., Stein, M. K., & Baxter, J. (1991). Where subject knowledge matters. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 2, pp. 87–113). London: JAI Press Inc.
Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, 90(3), 174-184.
Leron, U. and Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly*, 102, 227-242.
Lo, J., Grant, T., & Flowers, J. (2004). Developing mathematics justification: The case of prospective elementary school teachers. In D. McDougall & J. Ross (Eds.), *Proceedings of the twenty-sixth annual meeting of the north American chapter of the international group for the psychology of mathematics education* (Vol. 3, pp. 1159–1166). Toronto: University of Toronto.
Porteous, K. (1986). Children's appreciation of the significance of proof. In *Proceedings of the Tenth International Conference for the Psychology of Mathematics Education*, 392-397. London, England.
Martin, J. R. (1970). *Explaining, understanding, and teaching*. New York: McGraw-Hill.
Mejia-Ramos, J. P. & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna & M. de Villiers (Eds.), *Proceedings of the ICMI Study 19 conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 88-93), Taipei, Taiwan.

Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. Presentation at the *14th Conference for Research in Undergraduate Mathematics Education.* Abstract downloaded from: http://sigmaa.maa.org/rume/crume2011/Preliminary_Reports.html.

Rowland, T. (2001). 'Generic proofs in number theory.' In S. Campbell and R. Zazkis (Eds.) *Learning and teaching number theory: Research in cognition and instruction*. (pp. 157-184). Westport, CT: Ablex Publishing.

Schoenfeld, A. H. (2010). How and why do teachers explain things the way they do? In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 83–106). New York: Springer.

Speer, N., Smith III, J., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, 29, 99-114.

Thanheiser, E. (2009). Preservice elementary school teachers' conceptions of multidigit whole numbers. *Journal for Research in Mathematics Education, 40*, 252–281.

Thanheiser, E. (2010). Investigating further preservice teachers' conceptions of multidigit whole numbers: Refining a framework. *Educational Studies in Mathematics*, 75, 241-251.

Thompson, P. W., & Thompson, A. G. (1994). Talking about rates conceptually, part I: A teacher's struggle. *Journal for Research in Mathematics Education, 25*, 279–303.

Thompson, A. G., & Thompson, P. W. (1996). Talking about rates conceptually, part II: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 27*, 2–24.

Thurston, W.P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, 30, 161-177.

Weber, K. (2004). Traditional instruction in advanced mathematics classrooms: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, 23(2), 115-133.

Weber, K. & Alcock, L. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D. Stylianou, M. Blanton, & E. Knuth (eds.) *Teaching and learning proof across the grades: A K-16 perspective*. Routledge: New York.

APPENDIX: Materials used in this study

MASTER PROOF

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □

M1

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. To show f is injective, we must show that $f(x_1) \neq f(x_2)$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. It follows that f is injective. □

M2

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that

$$f'(x_3) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then

$$f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0.$$

Therefore $f(x_2) \neq f(x_1)$. □

M3

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. As $f(x_2) - f(x_1) \neq 0$, it follows that $f(x_2) \neq f(x_1)$. □

M4

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, f is a real valued function, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □

M5

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$, so $x_2 - x_1 = \frac{f(x_2)-f(x_1)}{f'(x_3)}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □

Communicating assessment criteria is not sufficient for influencing students' approaches to assessment tasks – Perspectives from a Differential Equations Class

Dann G. Mallet
Queensland University of Technology
dg.mallet@qut.edu.au

Jennifer Flegg
Queensland University of Technology, and
Oxford Centre for Collaborative Applied Mathematics, University of Oxford
j.flegg@qut.edu.au

This report presents the findings of an exploratory study into the perceptions held by students regarding the use of criterion-referenced assessment in an undergraduate differential equations class. Students in the class were largely unaware of the concept of criterion referencing and of the various interpretations that this concept has among mathematics educators. Our primary goal was to investigate whether explicitly presenting assessment criteria to students was useful to them and guided them in responding to assessment tasks. Quantitative data and qualitative feedback from students indicates that while students found the criteria easy to understand and useful in informing them as to how they would be graded, the manner in which they actually approached the assessment activity was not altered as a result of the use of explicitly communicated grading criteria.

Key words: differential equations, assessment experiment, criterion-referenced assessment

Introduction

Criterion-referenced assessment (CRA) is assessment that is constructed with the intent to measure student performance that can be explained with reference to clearly delineated learning tasks (Linn & Gronlund, 2000). It involves identifying the extent of a learner's achievement of predetermined goals or criteria and fundamentally involves assessing a student without reference to the performance of others (Brown, 1988; Harvey, 2004; TEDI, 2006). When CRA is used it requires an underlying set of course learning outcomes, an assessment program designed to gather information about a student's performance in relation to those learning outcomes, and importantly the communication of criteria and standards (a two dimensional view) which inform students how they will be judged and to provide directions for assessors.

Generally speaking, Australian Universities impose or very strongly encourage the use of criterion referenced or standards based assessment in the courses that they offer. At the authors' home institution this is no different with the University's Manual of Policies and Procedures stating that the University "has adopted a criterion-referenced approach to assessment where assessment is based on pre-determined and clearly articulated criteria and associated standards of knowledge, skills, competencies and/or capabilities" (QUT, 2011). Furthermore, the policy states that assessment is "clearly communicated to students" and used as "a strategy to support student learning" (QUT, 2011). In essence, the implication is that a particular method of assessment, CRA, is imposed on all teaching academics so as to ensure students are aware of how they are being assessed and because it is useful in supporting students in their learning. In this study, we challenge this perception by appealing to the thoughts of students themselves.

Until recently, the directive to employ CRA has been largely ignored in the context of many quantitative courses such as those in mathematics and the sciences. Lecturers regularly justify this; claiming that quantitative studies involve assessment responses that are either right or wrong and that right/wrong are sufficient. In this study we report on the successful implementation of elements of criterion referenced assessment into a Differential Equations course that goes beyond simple “right-wrong” criteria while maintaining the mathematical integrity of the assessment program. Specifically, we unpack the usual collection of “right” and “wrong” judgments and collect them into groups related to the learning outcomes of the course, thereby providing assessment criteria. We present findings based on quantitative and qualitative feedback from students regarding their perceptions of criterion referencing and how it is used in guiding their learning throughout the course.

It is important to place this study in context by comparing the assessment experiment with the methods previously used to assess students in the course. Over approximately the past 10 years, the course has been taught by four different lecturers, however the assessment strategy has essentially been to employ 1-2 assignments (problem solving tasks with a 2-4 week completion timeframe) and a mid-semester and final examination. These tasks generally contribute 30-40% (assignment) and 60-70% (examination) of the final grade for the course, respectively. Assessment of students on all of these tasks has been carried out using what we refer to as the “traditional method” for mathematics assessment and *not* criterion-referenced assessment. That is, the academic responsible for assessment writes an examination or assignment, along with his or her own set of “correct” solutions. The set of correct solutions is annotated with points or marks throughout the solutions where points correspond with reaching certain points in the solution process. Assessment using the traditional method involves making judgments as to whether a student is right or wrong at various points in a solution procedure and makes no explicit reference to the learning outcomes that the academic intends students to obtain as a result of undertaking the course of study.

In the assessment experiment reported on in this paper, we have attempted to maintain the previously employed assessment program as much as possible. In particular, we maintained progressive, non-examination assessment of 40% and used mid-semester and final examination contributing 60% of the students’ final grades. However, we implemented a criterion-referenced method of grading students in the assignment tasks completed during semester. This involved presenting students with a set of criteria and definition of standards in addition to the actual problems to be solved. Rather than simply implying that students would be marked right or wrong up to some number of points as is the case in the tradition method, students were provided with details of exactly how responses to the mathematical problems would be graded and how translation between the mathematics and the standards and criteria would be carried out.

Our goals in conducting this small-scale experiment fall into two main areas: to gauge students’ perceptions regarding criterion referenced assessment and its usefulness, and to a lesser extent, evaluating the motivation for effecting culture change among mathematics academics. With regard to students’ perceptions, we investigated how students viewed the understandability and the usefulness of criterion referencing and how they employed the additional information provided to them via the criteria and standards definitions in directing their learning and assessment responses. Implicitly, we believe that such an investigation and its results can then be used to effect culture change among mathematics teachers at universities by changing the way they view criterion referenced assessment, taking CRA from a directive imposed by administrators to a useful tool for mathematics learning.

This paper presents the analysis and implications obtained from a small-scale, mixed methods study of the use of criterion-referenced assessment in an undergraduate differential equations class. While the study has focused on a single course, the authors expect that the findings of the study could generally be carried over to other similar courses in applied mathematics, particularly in the Australian context. A similar study is currently being undertaken in a more advanced partial differential equations course and a comparison of findings in the two (albeit somewhat similar) contexts is currently under preparation.

Literature Review

While there is extensive practical experience and significant literature relating to the use of criterion-referenced assessment for mathematics at the school level, literature that describes the use of CRA in university level mathematics classrooms is close to nonexistent. Furthermore, at the time of undertaking this research, the authors were unable to find any published research discussing student perceptions regarding CRA and its impact on their learning process.

Niss (1998, in Pegg 2003, p.228) notes that mathematics assessment identifies and appraises the knowledge, insight, understanding, skill and performance of a student. Pegg however points out that this is not in fact the reality of assessment in mathematics and that rather, it is most often concerned with reproduction of facts and computational skills or algorithms (Pegg 2003). It is our contention that this is how previous years' assessment programs for the course under investigation have been presented to students. In the assessment experiment discussed in this report, we attempt to explicitly link the subtasks of the assessment activities with the learning outcomes of the course, which include such concepts as knowledge, insight and understanding in addition to skills. In this way we believe that our assessment becomes more of an educational tool for students than it has been in previous versions of the course, and that it allows for a more "constructive alignment" (in the sense of Biggs, 1996) of the content, pedagogy and assessment.

Criterion referenced assessment involves determining the extent to which a learner achieves certain predetermined goals or criteria, importantly, without reference to the performance of others (Brown, 1988; Harvey, 2004; TEDI, 2006). The implementation of CRA involves the design or statement of a set of learning outcomes for a course, design of a program of assessment to obtain information about a student's performance in relation to the learning outcomes, and the presentation of a criteria set and definition of standards which serves to both inform students how their performance will be judged and to provide directions for assessors.

Pegg (2003) notes that while the movement towards assessment based on outcomes and standards (rather than individual comparison) did initially have some basis in research regarding student learning, the links remain tenuous. As such, there is debate among teachers and academics alike as to whether the claims regarding the benefits of criterion referenced assessment are supported by strong research. Through research such as that presented in this study, we attempt to provide a research base that advocates the benefits and warns of the pitfalls of criterion-referenced assessment in the undergraduate mathematics classroom.

Conceptual Framework

In this study we carry out descriptive research related to questions around student perceptions and criterion referenced assessment. This descriptive research involves statistical and textual analysis/synthesis of data collected from a student population undertaking a course in differential equations in an attempt to understand student perceptions and provide guidance for academic staff in undertaking more useful assessment in mathematics courses. The context of an undergraduate differential equations course, described in the next section, was chosen due to its representativeness of typical applied mathematics courses and hence the potential for maximum transfer of findings across applied mathematics teaching and learning.

Context

The course that was used for the experiment described in this study was a second year undergraduate ordinary differential equations course. Content covered in the course included first and *n*th order linear equations, series solution methods, Laplace transform solutions, linear systems of differential equations, phase portraits, Bessel, Legendre and Cauchy-Euler equations and Fourier series solutions. While officially the prerequisite knowledge required for students to enter the course included advanced calculus or linear algebra (at the second year level), the actual requirement was only understanding first order ordinary differential equations as typically covered in first year level courses.

The cohort included 52 undergraduate students and 4 additional coursework postgraduate students (although the course content was second year undergraduate level). Teaching activities involved 3 hours per week of lectures presented in one 2 hour block and one 1 hour block to the entire groups as well as 1 hour per week of smaller group "tutorial" sessions with additional teaching assistance. The course ran for 13 weeks with a one-week mid-term break.

The official course outline lists the following learning outcomes for the differential equations course:

1. Engage your critical thinking skills to understand the principles of and develop theoretical knowledge regarding differential equations.
2. Draw on a range of your thinking skills to identify, define and solve real world and purely mathematical problems using existing knowledge and knowledge developed in this unit.
3. Communicate your theoretical understanding and problem solving attempts in methods appropriate to the context of this unit.
4. Demonstrate independence and self-reliance in retrieving and evaluating relevant information and in advancing your learning.

These rather broad objectives can be summarized as an intention to facilitate students developing critical thinking skills and theoretical knowledge, retrieving and evaluation relevant information, developing ability to identify, define and solve problems, and communicate results.

The assessment package included a 30% end semester exam, a 30% in semester exam (week 10), 2 problem solving tasks (weeks 4 and 8) totaling 30% and 2 short multiple choice quizzes (weeks 2 and 5) contributing 10% to the student's grade. It is the problem solving tasks, worth 15% each, that are specifically of interest in this study as these were the items assessed using explicitly communicated criterion referenced assessment. In particular, the marks allocated using a "traditional method" of assessment were analyzed and grouped into categories related to the learning outcomes of the course. The standards for each criterion were then determined by weighting with regard to the marks achieved in each category. We refer to this method of CRA as the "frequency-based standard allocation" and note for the reader's reference that taxonomy of standard allocations in applied mathematics CRA is the topic of other research (in preparation) by the authors.

Learning Outcome	Criterion	7 HD	6 D	5 C	4 P	3 MF	2 F	1 LF
2,4	*[20%] Correct identification of equation types and appropriate solution procedures*	Always or nearly always	Occasional errors	More than half time	About half the time	Less than half the time	Rarely	Never
1,2,4	*[70%] Apply existing and newly developed knowledge/skills to correctly arrive at solutions/results*	Consistent, effective and efficient throughout	Consistent and effective in most questions	Effective more than half the time	Around half the time	Less than half the time	Rarely correctly applied	Never or very rarely applies knowledge correctly
3	*[10%] Communication of mathematics/problem solving attempts*	Generally clear, logical	Minor/few errors and/or inconsistencies	Some errors and inconsistencies	Regular errors and inconsistencies	Many errors and logical inconsistencies	Rarely clear or logical	Meaningless or no attempt at communicating

Mapping of marks to questions and criteria Criteria	Q1	Q2	Q3	Q4	Q5
Correct identification of equation types and appropriate solution procedures	2	1	1	1	1
Apply existing and newly developed knowledge/skills to correctly arrive at solutions/results	4	3	5	4	5
Communication of mathematics/problem solving attempts	1	0.5	0.5	0.5	0.5

Figure 1: Criteria and standards along with a mapping of marks to criteria for one of the two problem solving tasks assessed using CRA in this study.

To maximize the feedback provided to students, and hence maximize support of student learning, students were provided with both the traditional feedback of "points" (with ticks and crosses) annotated upon their submissions and an annotated version of the criteria map shown in Figure 1. To fully elucidate the concept of frequency-based standards allocation employed in this study, we provide an example with reference to Figure 1. At the university where this study was undertaken a grade of 7 generally corresponds with a score of 85%-100% while a grade of 6 with a score of 75%-84%. Consider only criterion 1 and suppose a student scored 1 out of 2 points for question 1, and then 1 out of 1 point for each of the remaining questions. This gives 5/6 points or 83.3% for criterion 1 over the entire assessment item. Using the frequency-based standard allocation method this student would score a grade of 6 for criterion 1 on the problem solving task. The idea here is that students are providing not only with the tradition feedback of points, ticks and crosses, but also an explicit mapping of these ticks and crosses to criteria related to the learning outcomes of the course, with a view to providing them with deeper information regarding where they are progressing well and where they are struggling in the course.

Methods

We have used two primary data sources, one quantitative and one qualitative, in an attempt to address our research goals regarding student perceptions of criterion referenced assessment. The quantitative source involved a Likert scale survey while the qualitative tool comprised two questions to which students were requested to respond in free-text format. For maximum flexibility, both data collection tools were deployed online using SurveyMonkey. In order to gauge the full impact of the assessment experiment, surveys were conducted at the end of the course of study, following the provision of feedback to students on all criterion-referenced items and also following the post mid-semester exam feedback sessions. All 56 enrolled students were given the opportunity to take part in the survey and a 50% response rate was achieved at survey closure.

The first data collection tool was a 10-item survey using a 5 level Likert scale. Students were presented with the 10 items and asked to choose which response – strongly agree, agree, neutral, disagree or strongly disagree – best described their feeling regarding each statement in turn. The data collected is presented in full in Table 1. Numerical and statistical analyses of the Likert-survey were conducted with findings presented in the remaining two sections of this paper.

The second data collection tool was a survey allowing free-text responses on two questions of interest. The first question given to responders was

"*I would say that the impact of having an assessment task being marked by criterion referenced assessment on my approach to learning was …*"

while the second was

"*I see the educational benefits of using criterion references assessment as: …*".

The intention of the first question being to elucidate the student's perception of how CRA impacted on their own learning and their approach to assessment items completed *after* the problem solving tasks assessed using CRA. The second question was intended to obtain the student's wider view regarding the benefits of CRA. Textual analysis and synthesis was carried out on the free-text responses and again, findings and discussions are presented in the remaining sections of this work.

Results

The quantitative data collected via the first of the student surveys is presented in summary form in Table 1. The data indicates that while students found assessment criteria easy to understand and useful in informing them as to how they would be graded (items 1 and 10, 5 and 8), it did not alter the way they actually approached the assessment activity (item 2). Interestingly, on the whole it did not seem that students felt strongly that CRA provided more useful feedback to them than the traditional method in terms of preparing for future assessment (item 9). There was a similar almost uniform spread of responses regarding whether students found CRA useful at all (item 6).

	SA	**A**	**N**	**D**	**SD**
I feel that the details of the criterion referenced assessment (CRA) guidelines were made clear to me early in the semester.	**42.9% (12)**	39.3% (11)	14.3% (4)	3.6% (1)	0.0% (0)
I found that the way I approached completing the assessment task was different, given that I had the CRA sheet describing exactly how I would be graded.	14.3% (4)	14.3% (4)	10.7% (3)	25.0% (7)	**35.7% (10)**
Being assessed with a CRA sheet seems to me to be the best way that mathematics assessment can be graded.	3.6% (1)	17.9% (5)	**39.3% (11)**	14.3% (4)	25.0% (7)
I am so used to being assessed using CRA sheets that	0.0% (0)	21.4% (6)	**46.4% (13)**	17.9% (5)	14.3% (4)

being assessed in mathematics by this method did not concern me.					
The information provided by the CRA sheet made it clearer to me what was expected of me in order to get a particular grade.	17.9% (5)	**32.1% (9)**	25.0% (7)	17.9% (5)	7.1% (2)
I found the CRA sheets useful/helpful.	10.7% (3)	21.4% (6)	**32.1% (9)**	21.4% (6)	14.3% (4)
I believe I understand the educational benefits of CRA.	7.1% (2)	**42.9% (12)**	32.1% (9)	10.7% (3)	7.1% (2)
I feel to some extent that the CRA sheet demystified the way that marks are allocated in the assessment piece.	10.7% (3)	**42.9% (12)**	21.4% (6)	14.3% (4)	10.7% (3)
I feel that the CRA grade provided me with more feedback on how I had performed in the assessment task and where I could improve in the future than a mark out of a total does.	14.3% (4)	**25.0% (7)**	21.4% (6)	17.9% (5)	21.4% (6)
I found the categories on the CRA sheet (ie, communication, problem solving) easy to understand.	21.4% (6)	**67.9% (19)**	10.7% (3)	0.0% (0)	0.0% (0)

Table 1: Students were asked to read each item and select the response which best described how they feel about the statement. SA=strongly agree, A=agree, N=neutral, D=disagree and SD=strongly disagree.

Qualitative feedback from almost 100% of respondents indicated that in general the criteria provided were not used to determine how a student would approach individual questions or the assessment task as a whole. Interestingly, a similar percentage of students stated that they found CRA beneficial as it made the process of allocating scores by graders much clearer. A small percentage of students indicated that they did refer to the criteria sheets after the tasks were graded in order to get a different, higher level representation of where they had made errors in their responses.

Implications For Future Teaching Practice

The analysis of the data collected during this study indicates that while the concept and practice of CRA was clearly explained, and CRA sheets provided better guidance as to what was expected for different grade levels/marks, immediate and post-feedback learning approaches were not greatly altered. This research study has opened up new questions for future research. For example, we are now considering the impact on graders/academics and the usefulness they perceive in employing criterion referenced assessment.

With regard to application in the classroom in the future, both the qualitative and quantitative data indicate that students and graders alike, need to be explicitly informed exactly why they are provided with criteria and how they can be used to assist learning. Only 11 of 28 responded that they agreed in any way that CRA was more useful than traditional assessment for the purposes of preparing for future assessment items. Guiding students in their response attempts (showing them what the grader will deem to be “important”) and also aiding them in understanding the feedback they receive following the grading of their work are important benefits of CRA that should be communicated to students so that they may best use the feedback provided to them.

The actual construction of the criteria and standards is by no means straightforward. In the free-text responses, students indicated that in a general educational context they see CRA as providing better feedback, more guidance about how to approach a solution and an element of grading transparency. Clearly then, the process of constructing the criteria and standards is important, because these are where students gather this additional information and transparency. The criteria and standards must be carefully designed and worded so that they are exactly the types of judgments the grader is using while assessing students’ work. Academic staff need to be closely guided in the development of these elements of any criterion referenced assessment strategy.

Finally, we return to the point made in the introduction of this paper, namely that CRA is imposed on all teaching academics so as to ensure students are aware of how they are being assessed and because it is useful in supporting students in their learning. In this small-scale study, we have shed light on the fact that while CRA may be useful in raising student awareness about the assessment process, it is not sufficient in itself as the assessment method of choice to support students in the learning process. In fact, it may be more important to educate students regarding “how” to use feedback at all as a way to assist in their learning process, rather than to rely solely on the method itself.

References

Biggs, J. B. (1996) Enhancing teaching through constructive alignment, Higher Education, 32, pp. 347-364.

Brown, S. (1988) Criterion referenced assessment: what role for research? in Black, H.D. and Dockerell, W.D., New developments in educational assessment, British Journal of Educational Psychology, Monograph series no. 3, pp. 1-14.

Harvey, L. (2004). Analytic Quality Glossary http://www.qualityresearchinternational.com/glossary/#c. Accessed June 12, 2010.

Linn, R.L. and Gronlund, N.E. (2000). Measurement and Assessment in Teaching, (8th Ed). Upper Saddle River, NJ: Prentice Hall.

Pegg, J. (2003) Assessment in mathematics: A developmental approach, in Mathematical Cognition, J.M. Royer (Ed.), p. 227-259, Information Age Publishing.

Queensland University of Technology (QUT). (2011) QUT Manual of Policies and Procedures C/5.1 Assessment http://www.mopp.qut.edu.au/C/C_05_01.jsp#C_05_01.05.mdoc. Accessed March 24, 2011.

Teaching and Educational Development Institute (TEDI). (2006) Teaching and Learning Support: Criterion referenced assessment. http://www.tedi.uq.edu.au/teaching/assessment/cra.html. Accessed June 15, 2010.

an exploration of the transition to graduate school in mathematics

Sarah L. Marsh
University of Oklahoma
sarah.marsh@ou.edu

In recent years, researchers have given attention to the new mathematics graduate student as a mathematics instructor. In contrast, this study explores the academic side of the transition to graduate school in mathematics—the struggles students face, the expectations they must meet, and the strategies they use to deal with this new chapter in their academic experience. I will identify several resulting themes—Isolation vs. Community, Academic Relationships, Role of the Department, and Realizations of Self—from semi-structured interviews with mathematics graduate students designed to explore multiple aspects of the academic transition to graduate school. I will also use the social theory of legitimate peripheral participation (Herzig, 2002; Lave & Wenger, 1991) to discuss potential implications for graduate students.

Key words: graduate students, academic transition, semi-structured interview, legitimate peripheral participation

The academic transition from undergraduate to graduate school is undoubtedly a significant one in even the best of circumstances. During this transition, students often face new research expectations, increasingly abstract content, and an unfamiliar geographic location. Furthermore, in mathematics departments, many graduate students are also expected to teach or assist in undergraduate courses as graduate teaching assistants. Park (2004) noted that graduate teaching assistants can struggle with their new dual status as both learners and instructors, while Bozeman and Hughes (1999) referred to these experiences as an "abrupt change of status" (p. 347) for these students. Despite exemplary undergraduate records, students may have difficulty adjusting to their new environment or overcoming academic setbacks, such as insufficient prerequisite knowledge or an inability to meet professors' expectations.

Beyond diminishing students' self-esteem or their desire to complete their graduate study, these transitional stumbling blocks impact mathematics departments: Students' struggles with the transition to graduate mathematics may negatively affect program recruitment as admissions committees are less likely to admit applicants with similar backgrounds in the future. Retention is also impacted across the discipline as promising students may incorrectly assume they lack mathematical ability and leave the field forever. Finally, these struggles can affect the representation of women and minorities in such programs, as these groups are less likely to find the support structures they need to survive graduate school (Bozeman & Hughes, 1999).

To provide a foundation for future work in research on graduate students, I sought to establish a clear picture of what happens during the transition to graduate school in mathematics. This paper reports results from a larger, interview-based, qualitative study designed to explore the academic transition to graduate school in mathematics—the struggles students face, the expectations they must meet, and the strategies they use to deal with this new chapter in their academic experience. In particular, I started with the following exploratory research questions: What happens during the academic transition from undergraduate student to graduate student in mathematics? How do professors' expectations of new graduate students' mathematical knowledge affect students' success? How do new graduate students in mathematics adjust to the rigors of graduate school and/or compensate for prior knowledge deficiencies? How do attitudes, beliefs, and relationships play a role in the success of new graduate students in mathematics? I hope that this research will provide a more accurate picture of graduate student preparation for and experiences in graduate school in mathematics; then, we can work to modify resources for prospective and current graduate students accordingly to help make the transition as smooth as possible.

Background

Many academic transition points, such as the transition from secondary to tertiary mathematics, have been studied extensively as stakeholders have tried to narrow the achievement gaps among various groups of students. For instance, Selden (2005) discussed this transition to collegiate mathematics, noting that new college students must often reconceptualize ideas from previous mathematical training (such as the idea of a tangent line) in order to incorporate them into the new, demanding educational structure they have encountered. As another example, Kajander and Lovric (2005) detailed McMaster University's efforts to address this transition through surveys of students' mathematical backgrounds, course redesign, and provision of a departmental review manual to enable students' voluntary preparation for their mathematics courses. They noted that students' motivation, ability to delve beyond surface learning, and secondary school preparation in mathematics were all key to the transition process. Transferring the ideas from these two studies to the transition to graduate school in mathematics identifies several potentially relevant issues in this transition process: undergraduate preparation, ability to both reconceptualize prior knowledge and dig deeply into new mathematical material, and a "bridge" review process prior to graduate work. However, other factors may impact the transition to graduate school that are not found in this body of work.

Despite research on other transition points and recent work devoted to the new mathematics graduate student as a mathematics instructor (e.g., Border, 2009; Luo, Grady, & Bellows, 2001; Speer, Gutmann, & Murphy, 2005), little work has been done with other areas of the transition to graduate school, such as the impact of academic or personal issues on the transition experience. Related work has been done by Carlson (1999), who explored the mathematical beliefs and behaviors of "successful" graduate students; she found that persistence, high levels of confidence, and the presence of a mentor during key periods of mathematical development all played a role in these students' "success." However, while factors affecting retention and student success certainly impact students' experiences in the first months of a graduate program, they are not sufficient to define the transition to graduate school. More recently, Duffin and Simpson (2006) reported on interviews with Ph.D. students designed to explore the transition from undergraduate to graduate work in mathematics in the United Kingdom's educational system and concluded that both undergraduate and graduate education could be modified to smooth this transition for different types of learners. Despite this work, no clear picture of the transition to graduate school in the United States exists.

One theoretical lens through which to view research on graduate students—legitimate peripheral participation [LPP] in communities of practice (Lave & Wenger, 1991)—shows great promise. Lave and Wenger (1991) provide LPP as "a descriptor of engagement in social practice that entails learning as an integral constituent" (p. 35). They also describe LPP as "the process by which newcomers become part of a community of practice" (Lave & Wenger, 1991, p. 29). In 2002, Herzig applied this framework to a qualitative interview study examining persistence in graduate school in mathematics. In this case study of one mathematics department, Herzig interviewed both current students in the doctoral program and some who had left the program, as well as faculty members in this department, to investigate factors influencing doctoral student persistence and attrition. Herzig found that legitimate peripheral participation both in departmental life and in the field itself encouraged persistence in a doctoral mathematics program. Although this framework promises to enlighten research on graduate students more broadly, further work is needed to determine the phenomena involved in the graduate education of mathematics students. Thus, to further research in this young area, a more complete description of the transition experience is needed.

Research Design

This paper focuses on a subset of the data from a larger, exploratory, single-case study designed to delve into the in-depth meanings of one mathematics department's experiences with the transition to graduate school. To fully explore these experiences, I adapted elements of Herzig's (2002) design, conducting interviews with both graduate students and faculty members in one mathematics department. Semi-structured interviews with both graduate students and faculty members were centered around the research questions given above, with probing questions included as appropriate. The student interviews (median length 1:01:48), which are the focus of this paper, allowed me to ask specific questions about participants' experiences surrounding the transition to graduate school, while the faculty interviews (median length 1:03:00) were designed to provide a new perspective on the same aspects of the transition experience. Faculty interviewees came from a list of those who had recently held positions related to graduate students—such as Chair, Associate Chair, Graduate Director, or core course instructor—and who were willing to participate. This paper focuses on the graduate student interviews.

Graduate students in mathematics at a large, public, Midwestern research university were emailed and invited to participate in a brief online survey (eight questions; median length of completion = 0:02:50) to provide basic demographic information to aid in interviewee selection. All 13 interview participants were domestic graduate students who had taken core courses in the Ph.D. track at this university. [Fourteen students self-identified as meeting these criteria based on the survey. However, upon conducting interviews, it was determined that only 13 actually met study criteria; the other interview was omitted from the data presented here.] Of these 13 participants, six were male and seven were female; six had completed an undergraduate degree at this university; and year of program entry ranged from 2003 to 2010.

Interviews were audio-recorded and fully transcribed. I used an open coding procedure (Strauss & Corbin, 1990) to build a structure to this transition by merging preliminary codes to identify themes in the data. This structure is thus grounded in participants' views and in their words (Creswell, 2007), as demonstrated by the reliance on participant quotes as evidence throughout the presentation of results.

Results

In the interview excerpts presented below, student quotations are tagged with codes of the form S*yz*, where the "S" indicates a student interview and the two-digit number *yz* gives the identifier assigned to that participant. In this way, individual participants can be followed throughout the data presented. Within the student interview data, four main themes have emerged and are discussed below: Isolation vs. Community, Academic Relationships, Role of the Department, and Realizations of Self.

Isolation vs. Community

This section presents three ideas related to the contrast of isolation and community. Several student participants, including some who had completed their undergraduate studies at the same university, identified a feeling of isolation upon transitioning to graduate school. All participants mentioned the impact of the academic and social community of their fellow graduate students. The idea of competition also played an interesting part in the sense of community that students experienced during graduate school.

Unexpected isolation. Seven of the 13 students reported experiencing an unexpected sense of isolation upon their arrival at this graduate program. For some students, this isolation was social in nature: “Especially in the first year, I would go home every couple weekends…. That was probably the biggest change for me, just being away from home, not having all my friends right there, like in college” (S02). Other students made similar comments: “The place that I came from was very stable, very supportive. I had tons of friends. I had a church where I knew lots of people. They supported me…. And I didn’t have that here. I didn’t know anybody” (S13).

These students explained their isolation as a function of the loss of the proximity of support structures such as family, friends, or other social systems. Ironically, even three students who had done their undergraduate work at the same institution also reported experiencing some social isolation upon entering this new chapter in their academic experience.

However, other students experienced a sense of isolation that could be better classified as academic in nature. For instance, when discussing a struggle with coursework during the first year of graduate school, one student made the following comments:

> I didn’t expect to fall so far behind so fast. I wasn’t used to that. When that had happened in previous courses, like in undergraduate courses, there were always a few people that were right there with me. I could talk to people in the class… And in grad school, I didn’t see other people getting really, really confused. That’s not to say that there weren’t people getting really, really confused, of course. For some reason, I got the impression that everyone else was pretty up to speed with what was going on, so that made it more difficult. (S14)

This student went on to describe how these academically isolated experiences were followed by lower grades, lack of confidence, and an unpleasant feeling toward graduate mathematics. Together, academic and social isolation separated students from the sense of community they came to realize was instrumental in their ultimate success.

Role of community. Most student participants emphasized a specific role that community had played in their graduate experience. This role was often academic:

> I would recommend just developing relationships with your classmates to support each other with your classwork. It helps a lot… to just talk about the material with them, and you’ll learn a lot. Some of you might understand some things, and you’ll learn a lot just explaining those things to the other students… I’d recommend developing that as early as possible. (S04)

Another student went so far as to make the following claim: “I work in groups so much now… I rely on other people, and that’s been a significant change. I know I could not get through graduate school by myself” (S05). Certainly, for a student pursuing a Ph.D.—and thus, likely very talented in the field—such an admission speaks volumes toward the importance of community in the graduate school experience. Clearly, having peers with whom to work through assignments and other content was crucial to these students’ success.

However, community fulfilled another important function—that of emotional and social support during the rigors of graduate school. One student commented on experiences with this dual role of community:

> Study groups are great, because if you know something, you know it better by teaching it to someone else, or explaining the problem to someone else, and if they know something, then they can explain it to you…. And so, it helps to work out the kinks with other people.

> And you also get that camaraderie of 'Yay! We're all in this together.'… That was one of the biggest things, I think, that really helped me get through a lot of the courses was just having that support circle of the study group. (S11)

Clearly, the support of a community of peers had a huge impact on these students. However, despite emphasizing this support, students often used the language of competition to describe their graduate school experiences.

Competition. Without prompting, five students specifically mentioned "competition" or "competitiveness" when relating their initial experiences in graduate school. However, these students were only rarely referring to actual competition between students for grades, attention, or any other endowment of status. More often than not, they were referring instead to the enhanced level of mathematical quality exhibited by their new peers.

> There would be, even though we're not competing against each other in a real competitive way, the competition as far as where you are in the class, you kind of realized that that was going to be, not so much like it used to be…We were probably all surprised that we weren't at the top of the class when we were used to being at the top of the class. (S08)
>
> There was actually a spirit of competition for the first time when I started grad school. It was very subtle. It wasn't like everybody was fighting for grades or anything, but you can actually sit around and talk about, 'How did you do on this thing?' and 'What did you guys do?' and I think for the first time I was actually immersed in a mathematical culture…. For the first time I was actually sitting around having conversations with people about mathematics. (S01)

Clearly, these students had to adjust to the new "mathematical culture" they were experiencing, but they generally seemed to embrace this culture whole-heartedly.

However, for one participant, the idea of competition was all too real:

> The competition in the classes with the other students was different [than in undergraduate]… Where I came from… the students were more supportive. There was a big study group. I had six or seven people who were in most of my classes as we all went through math majors together. And when we sat down and did homework, there were a minimum of four or five of us working on any given homework set at a time. And here, it could happen that way, but the people who were involved in these study groups here were the cream of the crop from where they came from. So, egos kind of got in the way a little bit more…. It gets a little harder. The study groups didn't come together as well. It gets a little more competitive. Especially in one of the particular classes that I took that [first] fall, the professor… would rank everybody's grades based on what they had done, and whenever there was a natural break, those would be the A's, those would be the B's, and if there was a big natural break somewhere else, those were the C's, and so it really was a competition. If you wanted an A, you really had to beat everyone else. (S13)

Interestingly, S13 was the only participant in this study who was no longer pursuing a Ph.D. in this department, despite spending the first few semesters of graduate school enrolled in core Ph.D. coursework. While this evidence is anecdotal at best, it does seem to emphasize the importance of a sense of support amongst the mathematical community in graduate school.

Academic Relationships

Beyond relationships with their peers, students also encountered other relationships during their first few months in graduate school that had a great impact on their transition experiences. This section discusses the role of students' relationships with content and with their (instructing) professors and the ways in which these relationships affected students' overall transition experiences.

Relationship with content. Eleven of the 13 student interviewees discussed the increased difficulty they experienced in their initial coursework in graduate school. For instance, one student commented that "the amount of work expected from each of the courses is a little bit more than you might see in undergrad, either in the level of difficulty of the problems, or just the number of problems" (S11).

Furthermore, while all students acknowledged that they had expected graduate school to be "hard," many of them also emphasized that they could not have anticipated the exact nature of the difficulty with which they were presented. One of these students addressed this phenomenon: "I anticipated an advanced difficulty level. I don't guess you can really prepare yourself for that, though, until you actually do it" (S12).

For some, the new intellectual challenges presented by graduate school were exhilarating:

> I remember my first week being very exciting, and very positive, because I got to be immersed in math, and go to all these different seminars, and hear all these words I'd never heard before, and see people talking about all these interesting things. (S10)

However, for others, this difficulty was overwhelming or burdensome, and they felt that professors' expectations played a role. Student S08 mentioned lying awake at night, worrying because "sometimes I think [professors] thought that I should know more than I did."

Other students had to compensate for specific prior knowledge deficiencies: "There were a few topics that I had to look up myself and teach myself in other courses because the professor expected that we already knew that, and I didn't necessarily know that" (S11). But, sometimes lack of preparation spanned an entire course (most notably, topology):

> I had never seen any topology at all, so having to come in to a graduate level sequence in topology, that was kind of a big shocker. I tried to prep myself with the undergraduate topology…the semester before I knew I was going to start that sequence, and it was a lot to try and prep yourself for. So, I didn't really feel prepared in that area at all. (S13)

Students discussed compensating for under-preparation in content areas such as topology, linear algebra, and complex analysis. While most students found some way to cope—using the Internet, supplemental textbooks, peers, or instructors—with a lack of prior knowledge, these initial struggles with course content played an important role in students' impressions of their transitions to graduate school.

Relationship with professors. Content was not the only area in which graduate students had to forge new academic relationships, however. Often, experiences with content were compounded by relationships with (instructing) professors, either actual or perceived:

> I don't think it was the material at first that was difficult, it's just, it took me a while to adjust to new people. Back at my old university, I knew all the professors, and I knew what to expect…. I knew the style of my professors…. My biggest transition has been adjusting to the people. (S07)
>
> I got the impression that everyone else was pretty up to speed with what was going on, so that made it more difficult. It made it so that I felt like the only person who could help me would be the teacher, and some professors are less approachable than others. Let's just put it that way. (S14)

In the first quote, the transition to graduate school was compounded by the lack of familiarity with faculty at a new institution. However, in the second example, a sense of academic isolation among peers deteriorated into a true academic struggle when professors were not "approachable" to this student. And while many of the relationship difficulties centered on mastering the mathematical content, other students struggled with the loss of interpersonal contact with professors they had experienced as an undergraduate student:

> I was very used to working closely with my professors at {school}, because it was such a small school, which was quite a change when I came here… At {school}, I would be in their offices every afternoon talking to them about things, and here I don't go to my professors' offices all that often….We were kind of friends, so that was quite a change, too. (S05)

Nearly all students mentioned the importance of relationships with professors or the need for adapting to professors' teaching styles as a factor in their ultimate success in graduate school. These comments came in response to questions about professors' expectations, the ways in which courses had changed since undergraduate school, and even compensation strategies. Thus, professors seemed to play an integral role in multiple facets of the academic experience for these students.

Role of the Department

While professors certainly play a large role in setting the department's cultural standards and expectations, other aspects of the department also impact students' experiences in graduate school. The department's administrative roles—such as handling quality-of-student-life concerns, advising students, and assessing degree progress—also have a great impact on students' experiences with the transition to graduate school. While some of these roles are dependent upon the particular people or departmental policies involved, others, such as those discussed below, are general enough to have potential applications outside of this specific department.

Informing. All participants mentioned that they would have liked to have more information regarding issues ranging from tuition costs to teaching responsibilities, from paychecks to degree requirements, from health insurance to the amount of time and effort a doctoral program would require. Many students framed these comments in the context of course advising, saying things such as: "For the first two or three years, it just feels like you're kind of floating around. 'Well, everybody takes these courses, so I'm just going to take these courses'" (S01). These students failed to see how to build their own degree plan around departmental requirements. Many (especially more senior) students wished information regarding these requirements had been more readily available or had been emphasized during mandatory advising appointments.

The lack of information was also strikingly felt in the area of research expectations. Several students wished that the idea of research, an explanation of the procedures and types of work involved, and the length and depth required by the research process had been introduced earlier in their degree. One student summed up these feelings this way:

> If you're going to get a Ph.D., you need to want to do research. And, I don't know how that could be communicated, mostly because people coming out of an undergraduate [degree], they have no idea what it means to do research. That's an unknown. And I don't know that you could even instruct them as to what that is at that point. But, I think that needs to be communicated as quickly as possible to graduate students… so they can make the decision if they want to do it, or if they want to stop at the master's. (S06)

Other students were surprised at the active seminar culture in the department, their role in attending and presenting in these seminars, and the process of finding a research advisor. While systems of disseminating information to students vary widely among institutions, this was one area in which these participants saw room for improvement.

Mentoring. While students felt strongly about the importance of community during the graduate school experience, several wanted to take community a step further to combat the lack of information they received regarding the navigation of the graduate school experience. These students felt that a mentoring program for graduate students would have helped them with their transition to graduate school:

> I've heard people talk about this, but I haven't seen it in place…doing some sort of mentoring program where you match a graduate student with an older graduate student and possibly a professor so that maybe graduate students don't feel quite so alone when they start. (S05)

Another student emphasized that this mentoring arrangement need not be terribly formal: "Not that we would need a mentoring program, but it would be nice to know someone that you could ask, or that was kind of assigned to you to talk to, or something" (S03). By putting community to work in even this relaxed way, these students felt that their needs during the transition to graduate school could have been met more effectively.

Realizations of Self

In addition to commenting on aspects of the department, culture, and content that impacted their transition experiences, many participants also commented on their own perspective changes or personal growth through the graduate school experience. While these took many different forms, two key realizations are discussed below.

Dedication. Eleven students emphasized the importance of being committed to and persevering through their degree program:

> You have to decide you're going to do it. If you're wishy washy, it's over. I think that, just making the decision 'I'm going to do this, and I'm going to finish it regardless,' that's probably what you have to do, because it is going to be difficult, and if you're at anything thinking 'I don't want to put up with this kind of difficulty,' then you're not going to do it. (S06)

Overcoming the inevitable obstacles of difficult material, research struggles, and a new social and academic environment was quite a challenge for some, but their perseverance carried them through. One student described this perseverance this way: "Once I've committed to doing something, it's very important to me to follow through on it. So, if I encounter difficulty with something I've committed to, then I'm going to do my best to keep with it and succeed" (S04).

While dedication to the experience was important for all students, one student emphasized the primary role that perseverance played in ultimate success in graduate school. This student said, "If you really want to do this, perseverance is probably the biggest thing. I never would have been able to do this if I hadn't been resigned to figuring it out at some point" (S14). This student went on to say that "I didn't get by because of my knowledge of mathematics. That had absolutely nothing to do with it…. That was just kind of a by-product of the persevering" (S14). While mathematical aptitude is obviously important to obtaining an advanced degree in mathematics, this student felt that perseverance ultimately played a more important role.

Searching for a place. In addition to realizing the importance of sheer dedication to the degree, students also stumbled upon new perspectives on their relationships with academics and with life. When discussing their place within the "mathematical culture," students made comments such as this:

> In undergrad, I was the top one or two or three students in the class, and then I felt like I was middle of the pack or less in graduate school….When everybody is interested, and

> everybody's knowledgeable, it always happens that you kind of fall more towards the middle, or bottom. (S03)

Other students had to adjust their definitions of "academic achievement" or "success" in order to satisfy their intellectual drives or academic perceptions of themselves:

> My attitude has changed a little, but I still have the attitude that 'Well, I might not be able to do as well as they do, but I can do enough to get through this and really learn and do well'. (S08)

> I think of myself very much as an academic person….Throughout my entire life, I've always defined myself by my success in school…. I think that determination to keep that part of what defines me, to not let that go, helped me keep saying that I had to succeed…. The definition of doing well has gotten tweaked a little bit…. I stopped building everything on an A. To me, doing well didn't necessarily have to correspond to the grade anymore. But, I still need to do well. There's still that drive there. (S11)

For some students, graduate school meant defining a purpose or role in the greater world outside of academia:

> Graduate school felt like a waste of time at the beginning. It was like, 'I'm just doing math. Who cares? What does this matter? So that, figuring out my place in society, I felt like I wasn't quite as useful as I could be. (S05)

This student went on to realize that graduate school did not have to serve as a placeholder between an initial degree and an ultimate career, later stating the following: "I've realized that my life is happening right now. I don't have to wait for it to start. It is going on, and I can find satisfaction and purpose in whatever I'm doing" (S05).

This sense of purpose was echoed in other interviews. Over half of the student interviewees explained that their life outside of mathematics helped keep them grounded or allowed them to balance out the stresses of graduate work. One student seemed to take this idea to heart particularly strongly:

> The one thing that I have to keep coming back to is that this doesn't define who I am… I think a lot of people who are talented academically face that, like, 'Are we defined by our performance, or what? Who am I?' So, keeping it in perspective that this is not my entire life… that's helped a lot, because when it's not so all-consuming, I'm less stressed about it, and I'm able to learn and perform a lot better when I'm less stressed. (S12)

These students' realizations of themselves, their attitudes, and their level of commitment to their mathematical careers were clearly an important part of their graduate school experience and would have made a great impact on their transition to graduate school if discovered sooner in their academic careers.

Discussion

While many facets of the study of mathematics graduate students remain unexplored, this study took a qualitative look at the transition to graduate school in mathematics. The results section above used students' own words to describe the importance of community, academic relationships with new content and professors, the department's role in informing and mentoring, and students' realizations of aspects of themselves throughout the graduate school experience. In this section, we revisit the results in relation to previous work, with the theoretical lens of LPP, and in terms of participants' and other recommendations for departments and for students experiencing this transition.

Isolation vs. Community

Students clearly relied upon both the academic and the social support of their community of practice (here, the mathematics department at their university) to thrive in their initial graduate work. This is consistent with Herzig's (2002) findings that "participation in the academic and social communities of the department is a critical factor in doctoral student persistence" (p. 206) since initial success using this support would promote continued persistence in the program of study. Furthermore, Lave and Wenger's (1991) theory of the importance of LPP supports the idea that newcomers to a community of practice need to engage both with other members of the community and with tasks relevant to that community's work. In this way, both collaborating with peers on assignments or discussions of course material and socializing with those peers are a necessary part of the transition into the new academic community of graduate school.

Academic Relationships

For many students, an important part of the transition to graduate school was negotiating new relationships with professors and with increasingly difficult and abstract mathematical content. Beyond the horizontal sense of community mentioned above, students also need the vertical sense of community established when they are treated as "junior colleagues" (Herzig, 2002, p. 201) by faculty members. Spending time around these gatekeepers of the mathematical community helps students achieve both the mathematical ability and skill necessary to excel in the field as well as the social and cultural knowledge needed to feel like a true member of the mathematical community (Lave & Wenger, 1991).

However, social and cultural knowledge are of little use if a student cannot master the material needed to stay in a graduate program. Students often have to determine new ways to study and to process course material beyond those they used as undergraduate students (Duffin & Simpson, 2006). Furthermore, insufficient prior knowledge can cause the transition to be unbearable for many students. One remedy for this, suggested by S14, would be to have some kind of "Intro to Graduate Math" course or review manual during the summer prior to the beginning of graduate school. Such an idea has already been piloted at the secondary/tertiary transition level (Kajander & Lovric, 2005), although Kajander and Lovric (2005) noted that students' individual motivation to utilize review materials would play a key role in the success of such a program.

Role of the Department

Beyond what happens with mathematical content or course-related interactions, however, students must learn the "ins and outs" of succeeding in graduate school and in mathematical culture more broadly. Realistic expectations of the nature of the degree path (sets of departmental exams to be passed, paperwork required by various campus offices, average time to graduation) and of the research process can be gleaned while students participate in the community in legitimate, peripheral ways. However, many graduate programs in mathematics expect students to master prescribed content through coursework rather than encouraging them to "*think*, *act*, and *feel* as mathematicians do" (Herzig, 2004, p. 389). Thus, students in transition from undergraduate programs do not have access to these cultural forms of knowledge and can feel lost within the graduate system. Incorporating mentoring, which has been related to success in mathematics (Carlson, 1999), connects incoming students to the departmental culture via a vertical (advanced graduate student or professor) sense of community. These connections, combined with the sense of community among peers and the relationships with professors discussed above, could help students navigate the transition into graduate school more smoothly.

Realizations of Self

The students interviewed for this project strongly emphasized their dedication to, or persistence with, their graduate work. In 1999, Carlson also found that "persistence was the trait most frequently cited [by graduate student interviewees] as facilitating… mathematical success" (p. 244). Herzig (2002) noted that participation in "the academic and social communities of the department" (p. 206) played a key role in whether students persisted in their pursuit of a graduate degree. As noted above, the sole interviewee who had decided not to pursue a doctoral degree was one who had several negative things to say about the transition into the graduate program. While this could be viewed as an isolated case, we should also consider the role that transition experiences play in whether students persist toward the Ph.D.

Conclusion

As they progressed through graduate school, the graduate students interviewed for this study had to take on more and more tasks common to life as a mathematician. According to Lave and Wenger (1991),

> Moving toward full participation in practice involves not just a greater commitment of time, intensified effort, more and broader responsibilities within the community, and more difficult and risky tasks, but, more significantly, an increasing sense of identity as a master practitioner. (p. 111)

That is, as students become more involved in the teaching, research, and service activities common to a mathematician in an academic setting, they form their own sense of identity as a mathematician. While this identity formation happens over time as students participate in a community of practice, early experiences with the transition help shape students' view of the field and of "master practitioners" therein.

References

Border, L. L. B. (Series Ed.). (2009). *Studies in Graduate and Professional Student Development: Vol. 12. Research on graduate students as teachers of undergraduate mathematics.* Stillwater, OK: New Forums Press.

Bozeman, S. T., & Hughes, R. J. (1999). Smoothing the transition to graduate education. *Notices of the AMS, 46*, 347-348.

Carlson, M. P. (1999). The mathematical behavior of six successful mathematics graduate students: Influences leading to mathematical success. *Educational Studies in Mathematics, 40*, 237-258.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Duffin, J., & Simpson, A. (2006). The transition to independent graduate studies in mathematics. In F. Hitt, G. Harel, & A. Selden (Eds.), *Research in Collegiate Mathematics Education VI* (pp. 233-246). Washington, DC: American Mathematical Society.

Herzig, A. H. (2002). Where have all the students gone? Participation of doctoral students in authentic mathematical activity as a necessary condition for persistence toward the Ph.D. *Educational Studies in Mathematics, 50*, 177-212.

Herzig, A. H. (2004). 'Slaughtering this beautiful math': Graduate women choosing and leaving mathematics. *Gender and Education, 16*(3), 379-395.

Kajander, A., & Lovric, M. (2005). Transition from secondary to tertiary mathematics: McMaster University experience. *International Journal of Mathematical Education in Science and Technology, 36*, 149-160.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* New York: Cambridge.

Luo, J., Grady, M. L., & Bellows, L. H. (2001). Instructional issues for teaching assistants. *Innovative Higher Education, 25*, 209-230.

Park, C. (2004). The graduate teaching assistant (GTA): Lessons from North American experience. *Teaching in Higher Education, 9*(3), 349-361.

Selden, A. (2005). New developments and trends in tertiary mathematics education: Or, more of the same? *International Journal of Mathematical Education in Science and Technology, 36*, 131-147.

Speer, N. M., Gutmann, T., & Murphy, T. J. (2005). Mathematics teaching assistant preparation and development. *College Teaching, 53*(2), 75-80.

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

Students' Reinvention of Formal Definitions of Series and Pointwise Convergence

Jason Martin
Arizona State University
Jason.H.Martin@asu.edu

Michael Oehrtman
University of Northern Colorado
Michael.Oehrtman@unco.edu

Kyeong Hah Roh
Arizona State University
khroh@math.asu.edu

Craig Swinyard
University of Portland
Swinyard@up.edu

Catherine Hart-Weber
Arizona State University
Catherine.Hart@asu.edu

The purpose of this research was to gain insights into how calculus students might come to understand the formal definitions of sequence, series, and pointwise convergence. In this paper we discuss how one pair of students constructed a formal ε-N definition of series convergence following their prior reinvention of the formal definition of convergence for sequences. Their prior reinvention experience with sequences supported them to construct a series convergence definition and unpack its meaning. We then detail how their reinvention of a formal definition of series convergence aided them in the reinvention of pointwise convergence in the context of Taylor series. Focusing on particular x-values and describing the details of series convergence on vertical number lines helped students to transition to a definition of pointwise convergence. We claim that the instructional guidance provided to the students during the teaching experiment successfully supported them in meaningful reinvention of these definitions.

Keywords: Reinvention of Definitions, Series Convergence, Pointwise Convergence, Taylor Series

Introduction and Research Questions

How students come to reason coherently about the formal definition of series and pointwise convergence is a topic that has not be investigated in great detail. Research into how students develop an understanding of formal limit definitions has been largely restricted to either the limit of a function (Cottrill et al., 1996; Swinyard, in press) or the limit of a sequence (Cory & Garofalo, 2011; Oehrtman, Swinyard, Martin, Roh, & Hart-Weber, 2011; Roh, 2010). The general consensus among the few studies in this area is that calculus students have great difficulty reasoning coherently about formal definitions of limit (Bezuidenhout, 2001; Cornu, 1991; Tall, 1992; Williams, 1991). The majority of existing research literature on students' understanding of sequences and series concentrates on informal notions of convergence (Przenioslo, 2004) or the influence of visual reasoning or beliefs (Alcock & Simpson, 2004, 2005). Previous literature on pointwise convergence has been in the context of power series addressing student understanding of various convergence tests (Kung & Speer, 2010), the categorization of various conceptual images of convergence (Martin, 2009), the influence of visual images on student learning (Kidron & Zehavi, 2002), and the effects of metaphorical reasoning (Martin & Oehrtman, 2010). Two student volunteers from a second-semester calculus course participated in a teaching experiment with the goal that they reinvent the formal definitions of sequence, series, and pointwise convergence. For this paper we posed:

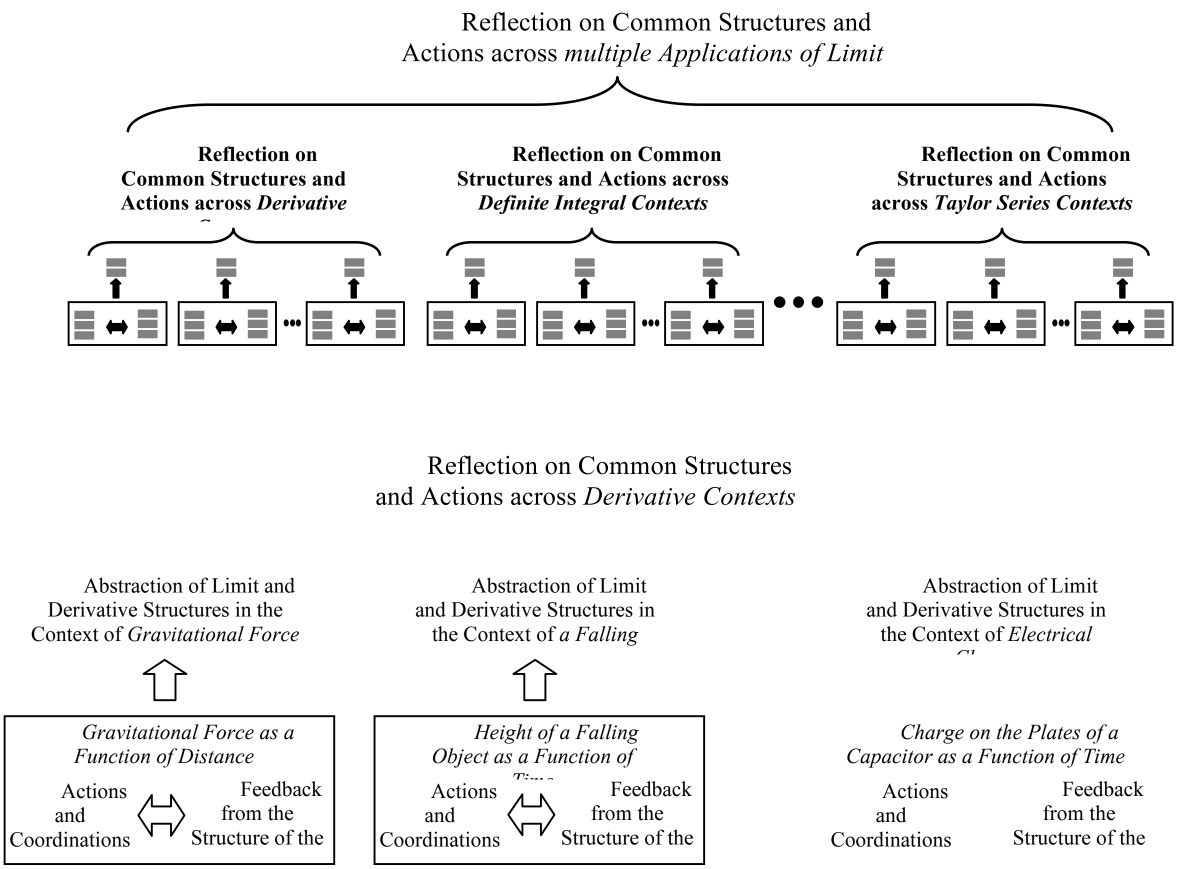
1. What are the challenges that students encountered during guided reinvention of the definitions for series and pointwise convergence? How did students resolve these difficulties?
2. What aspects of the students' definition of sequence convergence supported their reinvention of series convergence? What aspects of the students' definition of series convergence supported their reinvention of the definition of pointwise convergence?

Theoretical Perspective

To investigate our research questions, we adopted a *developmental research* design, described by Gravemeijer (1998) "to design instructional activities that (a) link up with the informal situated knowledge of the students, and (b) enable them to develop more sophisticated, abstract, formal knowledge, while (c) complying with the basic principle of intellectual autonomy" (p.279). Task design was supported by the *guided reinvention* heuristic, rooted in the theory of Realistic Mathematics Education (Freudenthal, 1973). Guided reinvention is described by Gravemeijer, Cobb, Bowers, and Whitenack, (2000) as "a process by which students formalize their informal understandings and intuitions" (p.237). The design of the instructional activities was inspired by the *proofs and refutations* design heuristic adapted by Larsen and Zandieh (2007) based on Lakatos' (1976) framework for historical mathematical discovery.

Oehrtman et al. (2011) further characterize reinvention in the context of students creating a formal definition as beginning with exploration of a rich set of examples and non-examples, then moving to an iterative refinement process that involves definition creation, checking examples, conflict acknowledgement, and discussion. During the multiple iterations of this process, students engage challenges that arise from their articulations of their definitions trying to formally capture their understandings of a particular concept. These challenges that students face during reinvention are identified by Oehrtman et al. (2011) as opportunities for learning. They identify two basic types of opportunities for learning: *problems* directly identified by students and *problematic issues* unbeknownst to the students but would be identified as problems by the mathematics community. Problems most commonly arise due to conflicts between the students' concept image and their currently stated definition. Furthermore, concerning the currently stated definition, experts might identify other problematic issues that have not been recognized by the students. Problematic issues may become identified problems by the students or may be indirectly resolved by the students while addressing another problem. Oehrtman et al. (2011) claim that it is the thoughtful resolution of identified *problems* that are most meaningful for students and therefore, support the formation of integral ideas that can lead to components of their definitions that remain stable throughout the remaining iterative refinement process.

Students were recruited from a second-semester calculus course covering topics such as sequences, series, and Taylor series, in addition to other topics typical to a second-semester calculus class using the textbook by Smith and Minton (2007). Furthermore, this course utilized Oehrman's (2008) approximation and error analysis framework as a coherent instructional approach to developing the concepts in calculus defined in terms of limits. This framework is based on an approximation metaphor identified by Oehrtman (2008, 2009) as including structures involving "estimates," "error," "accuracy," etc. which involve an unknown actual quantity and a known approximation. For each approximation there is an associated error and a bound on the error. The approximation is viewed as being accurate if the error is small. Actual student usage of the metaphor can be idiosyncratic (Martin & Oehrtman, 2010) but with repetitive usage of ideas of approximations, errors, and bounding errors to reinforce common limit structures within and across different limit contexts, students' use of the metaphor can become more systematized with a metaphorical structure that reflects the structure of formal limit definitions (Oehrtman, 2008). Conversely, a more systematic metaphor can encourage the abstraction of a common structure while engaging in multiple activities within a limit context and the results of such abstractions further support abstractions of common structures across different limit contexts that can provide a more coherent understanding of the role of limit throughout all of calculus and beyond (*Figure 9*). As a student's approximation schema becomes well organized and as more abstract understandings of limit emerge over a long period of time,



this instructional approach can lend itself to supporting students in constructing a more formal understanding of the limit concept through reflection upon common structures and actions performed on approximation tasks.

Methods

The authors conducted a six-day teaching experiment with two students, Megan and Belinda (pseudonyms), who had not received instruction on formal definitions for *sequence convergence*, *series convergence*, and *pointwise convergence*. The central objective of the teaching experiment was for the students to generate rigorous definitions for these three convergence definitions. The full teaching experiment was comprised of six, 90-120 minute sessions with this pair of students after the students had completed their class activities, including a chapter exam, on sequences, series, and Taylor series. The research reported here focuses on the evolution of the two students' definitions of series and pointwise convergence over the course of the 3rd through 5th

Figure 9. Layers of Abstraction

* From Layers of Abstraction: Theory and design for the instruction of limit concepts by Oehrtman, M.

sessions of the teaching experiment following the students' reinvention of a formal definition of sequence convergence.

During the first three days of the teaching experiment, the students engaged in the reinvention iterative refinement process in trying to produce a sequence convergence definition consistent with their concept image of sequence convergence (Oehrtman et al., 2011). Prior to the beginning of the teaching experiment, students from the calculus class engaged in a classroom activity to create an extensive list of graphical images depicting what they viewed as qualitatively different examples of sequences converging to 5 and sequences not converging to 5. Following the production of these graphs, in the teaching experiment Megan and Belinda (pseudonyms) were prompted by the facilitators, Craig and Jason, to construct a convergence definition for sequences by completing the statement, "A sequences converges to U as $n \to \infty$ provided…" Using their list of examples to test their definitions, over the next three sessions the students produced more than 23 definitions for sequence convergence before producing a definition for sequence convergence that they felt correctly captured the meaning of sequence convergence: "A sequence converges to U when $\forall \varepsilon > 0$ there exists some $N \geq 0$ $\forall n > N$ $|U - a_n| \leq \varepsilon$."

In leading up to producing their final definition Megan and Belinda either directly or indirectly engaged many challenges that provided them opportunity for learning through the thoughtful resolution of identified problems. As described in Oehrtman et al. (2011), some of these problems included:

1. How is our definition going to incorporate convergent sequences with "bad [random] early behavior?"
2. What is "infinitely close?"
3. A fixed error bound allows non-examples to be included as limits?
4. We have multiple uses of our "*n*" notation in our definition.

While wrestling with these problems, students had to mentally juggle ideas connected to indices, terms, errors, bounds on errors, universal and existential quantification, notational issues, graphical and formulaic representations, etc. It is after having thoughtfully resolved these issues in the production of their sequence convergence definition that the students were asked to produce convergence definitions for series and afterward, Taylor series, the two definitions that this report describes.

The teaching experiment activities on series began with Megan and Belinda producing and subsequently unpacking details of convergent series graphically. They were then asked to generate a definition by completing the statement, "A series converges when…" After these students had produced a definition of series convergence that they felt appropriately captured convergence in this context, the facilitators guided the discussion to issues of pointwise convergence. To address pointwise convergence, the students were asked to produce a graph of e^x with several approximating Taylor polynomials so that the graph would later be available for students to refer to when producing their Taylor series convergence definition. Later, the students were prompted to talk about what Taylor series were, and finally instructed to produce a definition for Taylor series convergence taking into account their Taylor series graph for e^x. The majority of each session consisted of students' iterative refinement of a definition and the unpacking of their intended meanings for individual elements within each definition.

As described in Oehrtman et al. (2011), there was five-member research team, with two researchers, Jason and Craig, serving as facilitators during the reinvention process. The team debriefed after each session and made adjustments to the next session's protocols as they made hypotheses about the progress of the reinvention. After the sessions were completed, the team transcribed video tapes and created content logs detailing each session with theoretical notes about student progress. A timeline of problems students were addressing was created so that we could attempt to determine both the origin and the resolution, if any, of a given problem.

Results

In this section we share some of the more illustrative examples of student interactions during the iterative refinement process of producing series and pointwise convergence definitions. Examples are chosen so as to illustrate 1) an overview of the reinvention process, 2) some problems students faced, whether directly identified by the students or not, 3) the process of coming to a solution to these problems, and 3) the influence of their prior reinvention experiences during the formation of these new definitions.

Series

In the 3rd session of the teaching experiment, the students initially drew a graph of an alternating series (top graph in *Figure **10***), but as they attempted to recall formulas for convergent series their production of additional series graphs stalled. When the facilitators prompted the students to not focus on finding formulas, the students compared sequence graphs to series graphs and expressed that series graphs "are harder to throw out there." After they stopped focusing on finding a formula they were able to produce a series graph increasing toward 7 (bottom graph in *Figure **10***). During this day, the students' initial and unprompted definition of series convergence to 7 was simply that a series converges when "the a_n's are going to 0 and s_n's are going to 7."

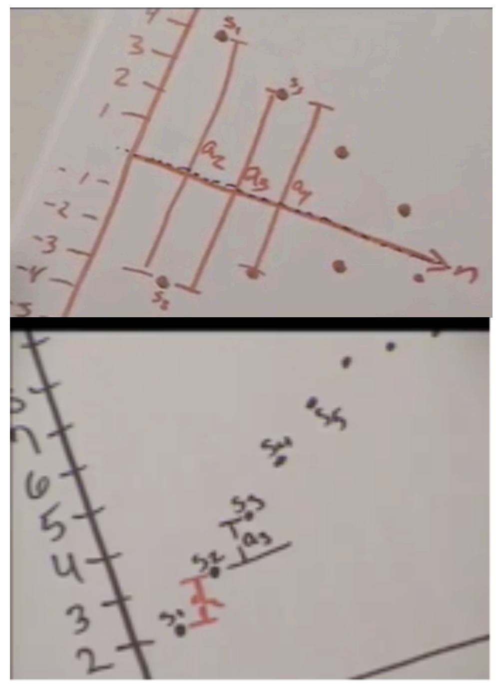


Figure 10. Examples of convergent series produced by the students.

During the 4th session, after being given the prompt to produce a series convergence definition, the students looked at their graphs of series convergence from the previous day and then immediately started making connections to their prior reinvention of sequence convergence.

Megan: Well, basically we could go with what we were talking about before. Only change it to, instead of the approximations, it's the partial sums, more or less. That would be a place to start.

Belinda: Yeah, 'cause I can see it [*looking at their graphs in Figure **10***]. It would be very similar to, to the sequences because you could still set the error bound within a certain, whatever range you want, any error bound,-

Megan: Uh-huh.

Belinda: -and then determine the point N where all the partial sums are within the error bound.

Megan: Uh-huh. Because graphically it looks more or less the same but instead of individual points, they are actually summations of the previous [terms], so instead of approximations, they're partial sums.
Belinda: Well, they're still technically approximations –
Megan: Yeah, yeah.
Belinda: -but they're, just instead of terms –
Megan: -but they're determined by summing the previous.

Belinda saw this activity as "very similar to" defining sequence convergence because she could still choose "any error bound" she wanted and that error bound would "determine" the N after which "all the partial sums are within that error bound." In these shorts statements, Belinda made references to N, ε (consistently attributed to "error bound"), universal quantification of ε, and the relationship of N to partial sums and ε's. Megan agreed, adding that "graphically" series "are more or less the same" to sequences. Furthermore, it should be noted that these students were evoking approximation language as they were reinterpreting the roles of terms, ε, and N in the context of series convergence.

Within 2 minutes they had written the definition, "A series converges to U when for all $\varepsilon > 0$ there exists a value N where all partial sums after N are within ε found by $|U - S_n| \leq \varepsilon$." For an hour, Megan and Belinda interpreted their definition graphically and formulaically using language, including approximation language, and notation from their prior reinvention experience. For instance, they reinterpreted elements, such as N, error bounds, and quantifiers from their prior definition of sequence convergence now as elements in a definition for series convergence. For example, while explaining their series definition they produced another convergent increasing series graph (*Figure* ***11***), and using this graph, they explained the effect of smaller error bounds on N:

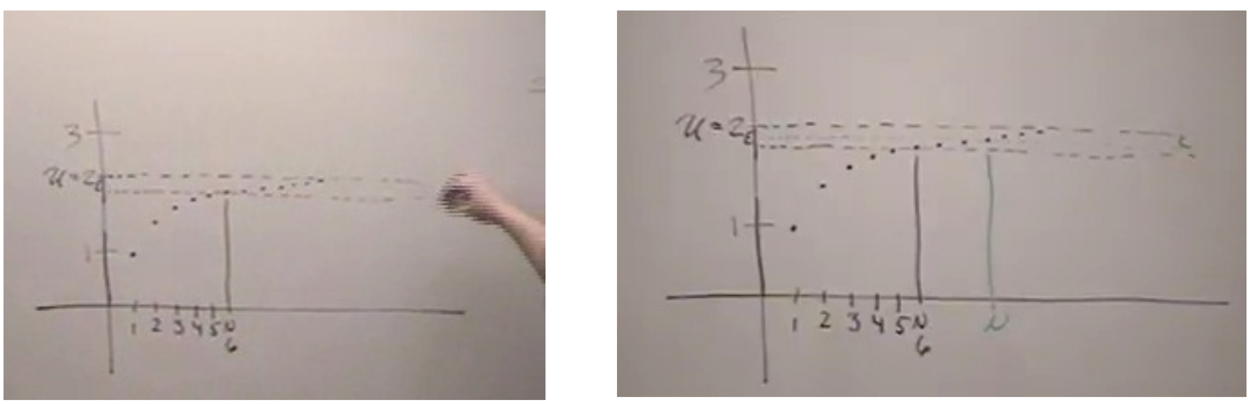


Figure 11. The effect of smaller error bounds on N

Belinda: So then I see how this one works because once we pick any sigma [*pointing to epsilon range depicted by the two horizontal dashed lines in the right graph in Figure* ***11***] we could move the sigma [epsilon] up or down. It wouldn't matter. For all of them, that exists we would have an N value [*see right graph in Figure* ***11***] where all the partial sums are within epsilon.

Megan: If we choose an epsilon [*drawing the a dashed green line appearing very faintly between the two horizontal dashed black lines found on the left graph in Figure **11***] here instead… that would make that our new *N* [*drawing the vertical line and the N in green in left graph in Figure **11***]. But after that *N*, these [*pointing to the partial sums to the right of her new N*] are all, still be within that new smaller [epsilon].

Furthermore, they recognized the need to replace their a_n notation for terms with the partial sum notation s_n. However, the students did not *just* change a_n to s_n, they considered each dot in the series graph as representative of partial sums, "You're adding a_1 to a_2 to a_3 to get each one of these dots on the graph of a series (as they illustrated in *Figure 12*)."

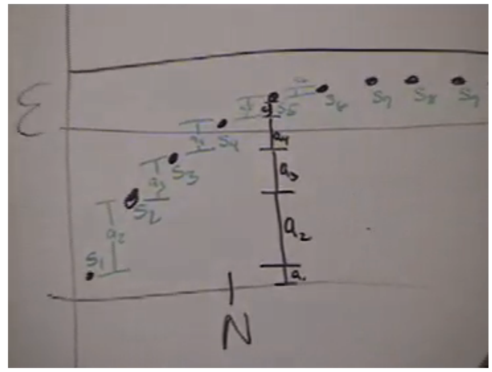


Figure 12. Dots are more than just dots

It should also be noted that they were directed to discuss their definition in the context of the divergent harmonic series. They initially recalled how their sequence convergence definition properly excluded sequences not converging to a given limit candidate when the sequence was in fact converging to something else. In light of this, they went on to explain how their series definition would exclude the harmonic series as being convergent to any particular given limit candidate because an *N* would not even exist for a given error bound.

After a few revisions, they constructed a definition for series convergence as follows: "A series converges to *U* when $\forall \varepsilon > 0$, there exists some *N* s.t. $\forall n \geq N$ $|U - S_n| \leq \varepsilon$."

Taylor Series

In initially discussing Taylor series during the 5th session, the students employed informal reasoning as they described various graphical attributes of Taylor polynomials approaching e^x. Following the first prompt, "What is a Taylor series?" students began talking about Taylor polynomials and their various approximation properties; such as how increasing the value for the index yields better approximations and the advantage of re-centering to produce better approximations with smaller values of the index for values of the independent variable away from the center of the series. Some of these discussions appear consistent with formal theory but their conceptions lying underneath their language was not consistent with formal theory.

The sameness of Taylor polynomials and the approximated function. During the later part of the discussion mentioned above, the following exchange occurred:

Megan: I think further out in the series it was the same as the function in question,
Belinda: Yeah, it actually became the function.
Megan: -for more of the graph. […] So like the first one was just tangent or something.
Belinda: Umm-hmm.
Megan: Or the second one was tangent and then other ones kind of started tracing along the graph-
Belilnda: Yeah.
Megan: -for further and further out. […]
Belinda: It was just becoming a way to actually create that function-

Megan: Umm-hmm.

Belinda: -the more n you went out. And depending upon how far you needed to be for that function, you could have less or more n to approximate it, to actually be that function.

Note how for Megan "the series" was the "same as" the approximated function and how the series "starts tracing along" "further and further out" as she moved from one Taylor polynomial to the next Taylor polynomial. For Belinda the series "actually became" and "actually created" the approximated function. Following these comments, Belinda went on to graph sin x and approximating Taylor polynomials. With this graph in front of them, Megan and Belinda continued to reiterate the sameness of the Taylor polynomials and the approximated function at more than just the center of the series. Belinda referred to Taylor polynomials as "actually looking like the sine graph." When asked what she meant, Megan injected that they "have the same values." Belinda agreed and added that the polynomials "fit the sine curve more and more" to the extent that one "couldn't distinguish the difference between the sine curve and Taylor [polynomials] at a certain point." Belinda illustrated what she meant by "certain point" by using 2π and concluded that after "some term" in the Taylor series, "the sine graph between 0 and 2π would be indistinguishable from the Taylor [polynomial]."

After the facilitators then moved the students to looking at their prior graph of e^x and illustrating errors at x=1 using their graphed Taylor polynomials (see *Figure 13*), Craig asked the students what would happen if "we kept taking approximations?" Belinda quickly concluded that "the approximation would equal e" and Megan agreed. Belinda then later clarified and said that "if you take out enough partial sums, then it, then the approximation will equal e^x at x=1." Craig then moved the students to talking about errors.

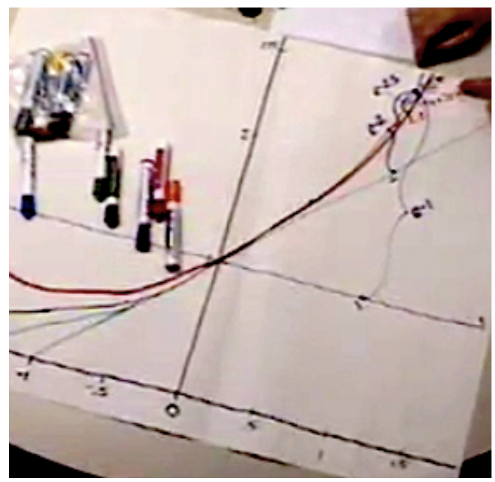


Figure 13. Errors Illustrated at x=1 on Taylor Series Graph for e^x

Craig: So if the approximation equaled e, what would the error be then?

Megan: Zero.

Belinda: It'd be zero.

Craig: Okay, and you're saying eventually your error will be literally zero. When do you think that would happen?

Megan: Based on where we're at right now, maybe another two approximations? If that?

Belinda: Yeah, it's possible. Two or three, I think. Maybe not even two or three. It could just be one more.

Even though their focus had moved to a particular x-value, they continued to employ informal reasoning that entailed Taylor polynomials and generating functions as being exactly the same after a certain number of terms had been added. Once they had used a Taylor series equation for e^x to find an explicit series for e they started to realize that a finite number of terms merely approximated e^x because the remaining terms not used in the approximation had value. Belinda was the first to come to this recognition after Craig highlighted the tail of the series shown in *Figure 14*. Belinda stated that "[the tail] is still technically not really zero" even though "it's very, very small." She then went on to say:

> Like up to here would be a good approximation [*underlining the partial sum for e*] but everything else, this total sum [*underlining the entire sum and putting an arrow on the right*] of everything would converge to *e* when *x* is 1. So this [*pointing to the partial sum for e*] just approximates *e*, but adding everything [*hand spread wide over the series as depicted in Figure 14*] to infinity converges to *e*.

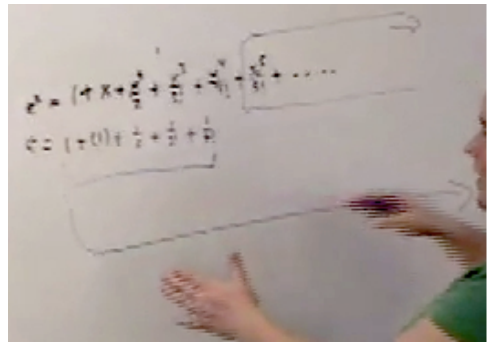


Figure 14. Tail of the Taylor series highlighted by Belinda as "not technically zero"

Belinda did not again make reference to Taylor polynomials being identical to the approximated function away from the center. When Megan later made a reference to a Taylor polynomial being "exactly" e^x at x=0.5, Belinda then proceeded to explain what it meant to "basically converge" at different values for x. Her explanations included references to polynomials getting "closer and closer" to e^x evaluated at corresponding x-values, errors getting "smaller" as more terms are added, and equality only achieved with the "total sum." Notions of polynomials now being exactly the same as the approximated function were missing from Belinda's explanations to Megan. Then, like Belinda, Megan's language started to embody the same ideas and Megan did not again make reference to Taylor polynomials as being identical to the approximated function away from the center.

Initial comparisons between series and Taylor series. When Belinda initially brought up "convergence," one of the facilitators asked, "What do you mean by converges to *e*?" Megan then began recalling those "definitions we wrote" for "sequence and series convergence." The students then made comparisons between series and Taylor series that mainly focused on the influence of the independent variable found in Taylor series but absent in series. For one, "series converge to a number but Taylor series converge to a function." Belinda even observed that Taylor series is "a bit more generalized" than series because "it doesn't have just one number that the Taylor series is converging to." Plus, according to Belinda, Taylor series is "a bit more flexible [than series] because you can either talk about the entire function or you can talk about specific points within the function." It was during these initial comparisons that a first attempt at a Taylor series convergence definition arose from Belinda: "If you add every partial sum going to infinity, then the Taylor series converges to whatever function it's approximating."

Confusion between the independent variable *x* and the index *n*. Following all the conversations depicted above, the students were given the prompt to complete the definition "$1+x+x^2/2+x^3/3!+x^4/4!+\cdots$ converges to e^x provided…" Immediately Belinda said, "Well, we could use the similar idea from sequence and series definitions." And then "given that this is a Taylor *series*, that [*pointing to their series convergence definition*] might be the better one to start with" (emphasis in original). So Megan began to write a Taylor series convergence definition but stopped after writing, "The Taylor series converges to $f(x)$." Meanwhile Belinda had been reading their prior series convergence definition and expressed concern over the role that N would play in a Taylor series convergence definition.

Belinda: I'm also thinking like we can't really use the *N*, like the cap *N* idea.
Megan: Yeah.
Belinda: Because there isn't really, I don't really see a cap *N* happening. You know what I mean?
Megan: Yeah.
Craig: What do you mean you don't see cap *N* happening?
Belinda: Well with a regular series it was easy to say at this cap, you know, if you set an error bound at-
Megan: There exists.
Belinda: -some cap *N*, […] All of the terms are going to be within that error bound. But this, the graphs look very different, so if we tried to use that, people aren't going to know what the heck is going on.

After these comments Belinda continued to reiterate how the graphs of series and Taylor series look very different. Not only in the approximations but in what they are converging to, "like a regular series graph isn't going to look like a func-, isn't going to be converging to a specific function." Belinda then turned around to their previously drawn Taylor series graph for e^x (see *Figure **15***):

Belinda: So if I used an *N*, like if I said, if I said an error bound, there's some *N* [*downward motion found in Figure **15***] where everything afterwards [*waving her hand to the right of where she did her downward gesture for N*] is going to be within that error bound, that's not going to make any sense.
Megan: It's not even going to work because-
Belinda: Because it's not going to work at all [*rightward gesture for error bound found in Figure **16***].
Craig: 'Cause it's-
Belinda: because right here I can set my error bound to be, like, that space apart [*repeats first picture in Figure **16***]. Then it's not going to work [*repeats similar gesture depicted in Figure **16***].

Based on her utterances and gestures, Belinda appeared to be directly overlaying the roles that ε and N play in two axis sequence and series graphs to their graph of Taylor series convergence for e^x. She correctly comes to the conclusion that in these roles, ε and N are "not going to work." Unfortunately, by applying ε and N in this way, she was neglecting the proper role of the index and the independent variable.

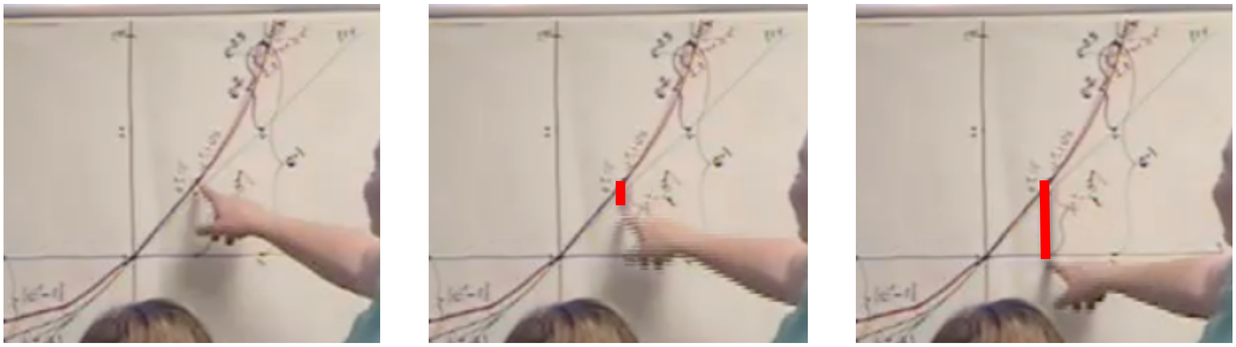


Figure 15. Belinda's Vertical Gesture of an "*N*" Value
*Red vertical "*N*" line added to illustrate what she was tracing.

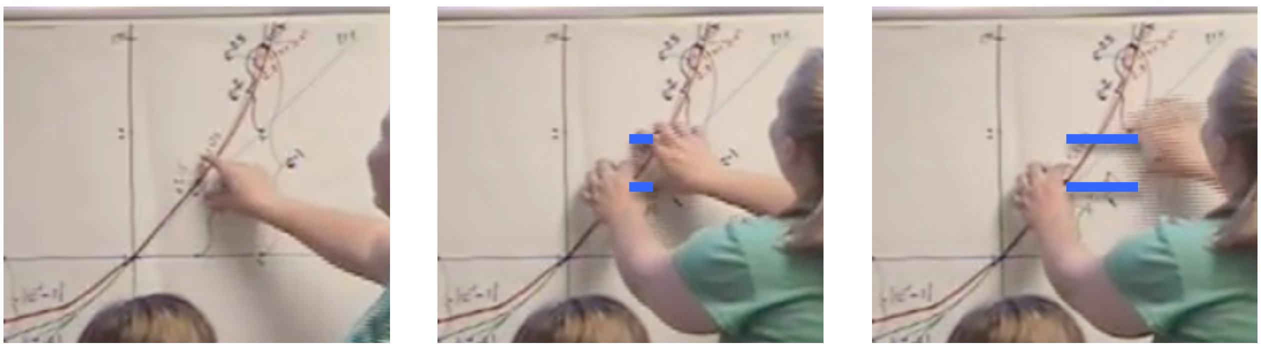


Figure 16. Belinda's Horizontal Gesture of an "Error Bound"
*Blue horizontal "error bound" lines added to illustrate what she was tracing.

Vertical number lines. In the calculus class students had depicted sequence and series graphs using not only the standard two axes graphs with axes for a_n and n, and S_n and n, respectively, they had also depicted sequence and series convergence on vertical number lines. *Figure 17* illustrates one such vertical number line for an increasing series converging to 5 drawn by Megan and Belinda. On the right of the number line, individual partial sums have been indicated using S_1, S_2, etc. and the terms composing the partial sum have been indicated as the difference between partial sums to the left of the number line using a_1, a_2, etc.

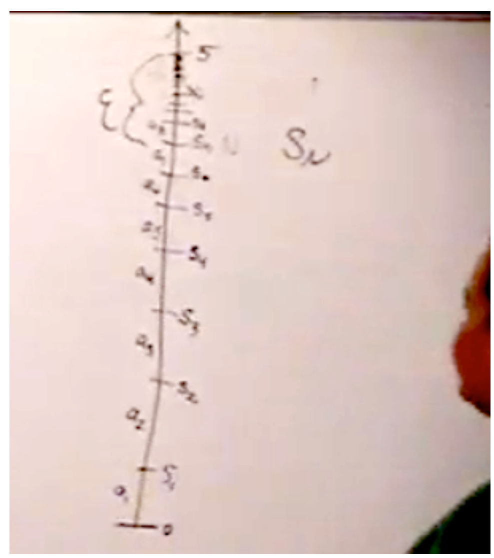


Figure 17. Vertical Number Line as Depicted by Megan and Belinda

Following a prompt from the facilitators to talk about their series definition for convergence in light of the vertical number line, the ε and N appearing in *Figure 17* (N appears very faintly in to the right of the number line) were added by Megan and Belinda while they were discussing components of their definition. Approximation terminology permeated their discussion as they consistently made references to approximations to 5 as the S_n partial sums, "errors" for approximations viewed as the difference between 5 and the corresponding approximation, "error bound" as ε, approximations "within" their error bound (the first such approximation denoted with index N), and how bounds on the error could be made small. From being given the prompt by Craig to producing a description of series convergence consistent with formal theory took the students around two minutes.

When the facilitators guided Megan and Belinda back to talking about Taylor series convergence, they then began to rearticulate their notion of convergence. During this rearticulation they both initially struggled in coordinating the roles of ε, N, and x. But then unprompted, they both drew graphs of functions resembling e^x with multiple vertical numbers lines illustrating multiple partial sum approximations for particular values of x. After another subsequent application of their series definition applied to one of these newly drawn vertical number lines appearing underneath the e^x like functions, Belinda suggested that they "expand that definition [*pointing to their series definition*] for all x-values that exist" because, as she would later say, "the Taylor series is essentially these number lines but there's just a whole lot of them."

Their final Taylor series convergence definitions. When they initially started to expand their definition, while reading both their series convergence and their prior start to a Taylor series convergence definitions, Megan suggests, "So we can say convergence to $f(x)$ for all x when for all epsilon greater than zero, dot, dot, dot." Even though Megan indicated putting the "for all x" at the beginning of their definition, Belinda suggested putting "for all x" at the end and the following definition was produced: "A Taylor series converges to $f(x)$ when $\forall \varepsilon > 0$ there exists some N such that $\forall n \geq N$ $\left|f(x) - S_n\right| \leq \varepsilon$ for all x."

After more probing by the facilitators using their previously drawn vertical number lines, the students clearly articulated that given a fixed error bound, N is different for different values for x and the same N that works for a specific value of x also works for all those values of x closer to the center but not the other way around. Apparently building off of this idea, instead of moving "for all x" forward in their definition, they suggest that their N in this definition is an N that is not found until "all the way up to infinity [*Belinda waving right hand to the right of the graph*]" and that this N works for all values of x. For the students this appeared to be some sort of idealized N that corresponded to x at infinity and if so, since an N for a specific value of x works for all those values of x closer to the center, this N at infinity would work for all values of x. But this N, as Belinda admitted caused a problem because it "may be impossible to find."

After the students admitted the potentiality impossibility of finding such an N, the facilitators brought the students back to their articulation of their definition that included the "for all x" at the beginning. When they reconsidered putting the "for all x" at the beginning, they immediately stated:

Megan: I think if we put this up here it would be restricting the N to be within epsilon of the x that we've chosen.
Belinda: I see what you mean. So like, depending upon what x-value we're at-
Megan: Yeah.
Belinda: -we would have a new N.
Megan: Yeah, and that would make it so we're not worrying about N not working further out.
Belinda: Yeah, try to find some N that's impossible to find.

For them, moving the "for all x" to the beginning of the definition resolved the problem they identified as trying to "find some N that's impossible to find" because their initial "N" had to work for all values of x. After some more discussion, which include them talking about S_n's dependence upon x, they eventually produced their final definition: "A Taylor series converges to $f(x)$ when $\forall x$, $\forall \varepsilon > 0$ there exists some N such that $\forall n \geq N$ $|f(x) - S_n(x)| \leq \varepsilon$."

It should be noted that following the production of this definition, the facilitators probed the students' understandings and the students described the roles for x, ε, N, and n consistent with formal theory. Even though they admitted that their last definition's placement of "for all x" causing consecutive universal quantifiers felt "goofy," they felt like their definition now best captured Taylor series convergence.

Conclusion and Discussion

It is remarkable that the students reinvented and unpacked the formal definition of series and pointwise convergence within such a short time. During the iterative refinement process for these definitions, the students faced multiple problems, whether the problems be problematic issues not directly identified by the student or problems directly identified by the student. One of the main problems that students faced during their reinvention activies of producing a series definition was the construction of series graphs. Problems were more numerous when the students were attempting to reinvent a Taylor series convergence definition. One of the more notable problematic issues was the students viewing the approximated function and Taylor polynomial as the same on some sort of growing interval of exactness that grew as the index of the Taylor polynomial increased. Some of the problems directly identified by the students after they were attempting to leverage their series convegence definition included how to handle the N in light of the independent variable and how to incorporate the "for all x" into their Taylor series convergence defintion.

We claim that the timely instructional guidance provided to the students during the teaching experiment successfully supported them engaging these challenges and their subsequent reinvention of these definitions. For example, the facilitators having students produce graphs of series and Taylor series convergence gave students a reference point for which they could refer to during the construction of their definitions. The guidance provided by the facilitors suggesting to the students that they produce series graphs without focusing on particular formulas freed the students from being constrained by attempting to recall specific series formulas and allowed them to produce more series graphs. The facilitators served as eventual conflict producers by having the students produce Taylor series formulas that led to their eventual recognition that the tail of the series had value and therefore, a Taylor polynomial and the approximated function in these cases could not be identical over an interval. The guidance to produce and unpack vertical number line graphs supported students in seeing the graphs of Taylor series as comprised of convergent series at each x-value where N is dependent upon x as well as ε. After students had wrestled with the problem of where to put the "for all x," the facilitators acted as solution providers and suggested that they reconsider their placement of "for all x" to best capture N's dependence upon x.

We also claim that the prior activity of defining sequence convergence became a means for supporting the students' definitions of series and Taylor series convergence as they recognized similarities between sequences, series, and Taylor series convergence in the context of their graphs and their definitions. For example, problems such as addressing the meaning of "infinitely close," the quantification of the error bound, and multiple uses of the "n" notation, which had all been issues during the reinvention of a sequence definition (see Oehrtman et al., 2011), were non-issues during their reinvention of series and Taylor series convergence definitions. Furthermore, when presented with the task of defining series convergence, they immediately saw similarities and started connecting elements from their sequence convergence definition to their forming series convergence defintion. Additionally, these connections were strong as evidenced by them being able to produce a formal series convergence definition in a matter of a couple minutes. Likewise, it is through the students' reflection on the components of their series convergence definition in the context of vertical number lines, the role of vertical number lines in the context of Taylor series graphs, and their coordination of the "for all x" with their prior series definition that they came to eventually produce a pointwise convergence definition for Taylor series consistent with formal theory.

Furthermore, the students' emerging approximation scheme greatly supported the students in meaningfully recognizing similarities between definitions and interpreting components within a definition. First, the approximation terminology that they had learned from class allowed them to meaningfully interpret the role of approximations, error, and error bounds within a limit context. For example, in the context of series, this was evidenced by references to partial sums as approximations, $|U - a_n|$'s as errors, and ε's as error bounds together with corresponding graphical interpretations. Second, the approximation terminology had been abstracted by the students in such a way as to allow them to use this terminology to identify and coordinate similarities between structural elements across different limit contexts. For example, this was evidenced by their references to approximations as terms of a sequence, partial sums of a series, and Taylor polynomials. During the creation of their series convergence definition, the approximation terminology helped bridge the gap between similar components of their sequence definition and their emerging series definition. Therefore, their emerging approximation scheme helped to organize their series definition in light of their sequence definition. Likewise, this type of interplay between elements of their series definition and their emerging Taylor series convergence definition using approximation terminology can be seen during their creation of their Taylor series convergence definition. In essence, this demonstrates how the approximation framework can support students in the highlest level of abstraction indicated in *Figure 9* and how this abstraction can further support students through reflection upon common structures and actions performed on relevant tasks in class and during reinvention.

As in Oehrtman et al. (2011), we acknowledge that individual students follow unique learning paths and that orchestrating the type of discussions needed for reinvention within an entire class will involve significant differences in what we have described here. Even so, this study demonstrates that reinvention of these definitions can occur and how it might occur. Furthermore, it demonstrates potential cognitive challenges, how challenges may be resolved, and how students can then leverage their resolution experiences to engage new challenges. We look forward to using these results to guide our future work to develop classroom activities for introductory analysis courses.

References

Alcock, L. J., & Simpson, A. P. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 57*(1), 1-32.

Alcock, L. J., & Simpson, A. P. (2005). Convergence of sequences and series 2: Interactions between nonvisual reasoning and the learners beliefs about their own role. *Educational Studies in Mathematics, 58*(1), 77-100.

Bezuidenhout, J. (2001). Limits and continuity: Some conceptions of first-year students. *International Journal of Mathematical Education in Science & Technology, 32*(4), 487-500.

Cornu, B. (1991). Limits. In D. O. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 153-166). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cory, B., & Garofalo, J. (2011). Using dynamic sketches to enhance preservice secondary mathematics teachers' understanding of limits of sequences. *Journal for Research in Mathematics Education*, *42*(1), 68-100.

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

Freudenthal, H. (1973). *Mathematics as an Educational Task*. Dordrecht, The Netherlands: Reidel.

Gravemeijer, K. (1998). Developmental research as a research method. In J. Kilpatrick & A. Sierpinska (Eds.), *Mathematics Education as a Research Domain: A Search for Identity (ICMI Study Publication)* (Book 2, pp. 277-297). Dordrecht, The Netherlands: Kluwer.

Gravemeijer, K., Cobb, P., Bowers, J., and Whitenack, J. (2000). Symbolizing, modeling and instructional design. In P. Cobb, E. Yackel and K. McClain (Eds.), *Symbolizing and Communicating in Mathematics Classrooms*, Erlbaum, Mahwah, NJ, pp. 225–273.

Kidron, I., & Zehavi, N. (2002). The role of animation in teaching the limit concept. *International Journal of Computer Algebra in Mathematics Education, 9*(3), 205-227.

Kung, D., & Speer, N. (2010). Do they really get it? Evaluating evidence of student understanding of power series. Paper presented at the 13th Annual Conference on Research in Undergraduate Mathematics Education, Raleigh, NC: North Carolina State University.

Lakatos, I. (1976). *Proofs and refutations*. Cambridge: Cambridge University Press.

Larsen, S., & Zandieh, M. (2007). Proofs and refutations in the undergraduate mathematics classroom. *Educational studies in mathematics*, *67*, 3, 205-216.

Martin (2009). *Expert Conceptualizations of the Convergence of Taylor Series Yesterday, Today, and Tomorrow*. Unpublished Dissertation, The University of Oklahoma, Norman, OK.

Martin, J., & Oehrtman, M. (2010). Strong metaphors for the concept of convergence of Taylor series. Paper presented at the 13th Annual Conference on Research in Undergraduate Mathematics Education, Raleigh, NC: North Carolina State University.

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: Mathematcial Association of America.

Oehrtman, M. (2009). Collapsing dimensions, physical limitations, and other student metaphors for limit concepts. *Journal for Research in Mathematics Education*, 40(4), 396-426.

Oehrtman, M., Swinyard, C., Martin, J., Roh, K., & Hart-Weber, C. (2011). From intuition to rigor: Calculus students' reinvention of the definition of sequence convergence. Paper presented at the 14th Annual Conference on Research in Undergraduate Mathematics Education, Portland, OR: Portland State University.

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. (2010). An empirical study of students' understanding of a logical structure in the definition of the limit of a sequence via the ε-strip activity. *Educational Studies in Mathematics*. *73*, 263-279.

Smith, R., & Minton, R. (2007). *Calculus: Early Transcendental Functions*. 3rd ed. New York: McGraw-Hill
Swinyard, C. (in press). Reinventing the Formal Definition of Limit: The Case of Amy & Mike. To appear in *Journal of Mathematical Behavior*.
Tall, D.O. (1992). The transition to advanced mathematical thinking: Functions, limits, infinity, and proof. In D.A. Grouws (Ed.) *Handbook of research on mathematics teaching and learning*. New York: Macmillan, 495-511.
Williams, S. R. (1991). Models of limit held by college calculus students. *Journal for Research in Mathematics Education, 22*(3), 219-236.

INQUIRY-BASED AND DIDACTIC INSTRUCTION IN A COMPUTER-ASSISTED CONTEXT

John C. Mayer[#], Rachel D. Cochran*, Jason S. Fulmore*, Thomas O. Ingram*, Laura R. Stansell[#], and William O. Bond[#]

University of Alabama at Birmingham (UAB)
[#]Department of Mathematics
*Center for Educational Accountability (CEA)

We compare the effect of incorporating inquiry-based sessions versus traditional lecture sessions, and a blend of the two approaches, in an elementary algebra course in which the pedagogy consistent among treatments is computer-assisted instruction. Our research hypothesis is that inquiry-based sessions benefit students significantly in terms of mathematical content knowledge, problem-solving, and communications. All students receive the same computer-assisted instruction component. Students are randomly assigned for the semester to one of three treatments (two inquiry-based meetings, two lecture meeting, or one of each, weekly). Measures, including pre- and post-tests with both open-ended and objective items, are described. Statistically significant differences have previously been observed in similar quasi-experimental studies of multiple sections of finite mathematics (Fall, 2008) and elementary algebra (Fall, 2009) with two treatments. Undergraduates, including many pre-service elementary teachers, who do not place into a credit-bearing mathematics course take this developmental algebra course.

Keywords. Elementary algebra, teaching experiment, computer-assisted instruction, inquiry-based instruction, didactic instruction.

One direction taken by course reform over the past few years has been the development of sophisticated computer-assisted instruction. This approach has been applied to large-enrollment service courses in mathematics, including algebra. Elementary algebra is typically taken by under-graduate students who do not place into a credit-bearing course. Traditionally, the goal of such a developmental algebra course has been to enhance students' "algebra skills," for example, dealing procedurally with rational numbers and expressions. Higher-order thinking may be largely absent. Alternately, one might focus on developing quantitative reasoning and communications skills, rather than, or in addition to, training to acquire a set of specific algebraic skills (Wiggins, 1989; Blais, 1988). Our position is that incorporating an inquiry-based component, either together with, or in place of, a didactic component, into a computer-assisted instructional environment may enhance student learning. Two previous studies in the literature bear this out (Mayer, 2009, 2010).

Fundamental Question

We compare three treatments in a quasi-experimental design: (GG) two weekly inquiry-based class meetings, (LL) two weekly lecture meetings, and (GL) one of each meeting weekly. The computer-assisted component is the same for all treatments. Our main hypothesis is that, of the three treatments, the one affording the most inquiry-based involvement to the students will differentially benefit the students in terms of mathematical content knowledge, reasoning and problem-solving ability, and communications. Secondarily, we expected no difference in student course grades among treatments. Typically, about 70% of students in this course earn an A, B, or C. Our hypothesis was supported in the areas of reasoning and problem-solving ability, communications, and course grades, but not supported in the area of mathematical content knowledge identified on the objective component of our pre/post-test.

Prior Research

Prior to the two most recent studies (Mayer, 2009, 2010), the methodology of simultaneously comparing different pedagogies within one semester, had few direct comparisons in the literature (Doorn, 2007). Some studies have compared different pedagogies over a longer time frame (Gautreau, 1997; Hoellwarth, 2005). The results of the quasi-experimental studies in (Mayer, 2009) of a finite mathematic course, and in (Mayer, 2010) of an elementary algebra course showed in both cases that students in the inquiry-based treatment did significantly better ($p<0.05$) comparing pre-test and post-test performance in the areas of problem identification, problem-solving, and explanation. Moreover, students, regardless of treatment, performed similarly (no statistically significant differences) when compared on the basis of course test scores. Outcomes of the two studies differed in gain in accuracy, pre- to post-test: in the finite mathematics study, there was no significant difference between treatments; in the elementary algebra study there was a significant difference between treatments in favor of the inquiry-based treatment. A limitation of both studies by Mayer was that accuracy was assessed on a small set of open-ended problems. The previous studies also did not test a blend of inquiry-based and traditional class meetings in a single treatment (Marrongelle, 2008).

Research Methodology

Our methodology is quasi-experimental in that it seeks to remove from consideration as many confounding factors as possible, and to assign treatment on as random a basis as possible, constrained only by students being able to choose the time slot in which they take the course. All students involved in the courses received identical computer-assisted instruction provided in a mathematics learning laboratory. 86% of the grade in the course was determined by evaluation in the computer-assisted context (online homework and supervised online quizzes and tests). The remaining 14% of the grade was determined by one of three pedagogical treatments, described below. Students registered for one of three time periods in the Fall 2010 semester schedule, a 9:00 AM, 10:00 AM or noon time slot, for three days a week (MWF), for their 50 minute class meetings and 50 minute required lab meeting. Students in each time slot were randomly assigned to one of the three treatments for the semester. Three instructors agreed to participate in the experiment. Each instructor taught in three time slots. In one slot the instructor administered the twice-weekly inquiry-based treatment, in another time slot, the twice weekly lecture treatment, and in a third time slot, the blended treatment. The three instructors consisted of a full professor, a regular full-time instructor, and a graduate student with prior teaching experience. All instructors had previous experience in both didactic and inquiry-based teaching, and in computer-assisted instruction. A graduate teaching assistant worked with each instructor in the inquiry-based meetings, and in evaluating written student work product from such meetings. Each instructor also met with each class in the mathematics computer lab. The computer lab meeting for all treatments occurred on Wednesday, and the class meetings on Monday and Friday. In the GL treatment, the lecture meeting was on Friday.

The three pedagogies to be compared are:

GG: two sessions weekly of inquiry-based group work (random, weekly changing, groups of four) without prior instruction, on problems intended to motivate the topics to be covered in computer-assisted instruction;

LL: two sessions weekly of traditional summary lecture with teacher-presented examples on the topics to be covered in computer-assisted instruction, and

GL: a blend of treatments GG and LL, with one weekly meeting traditional lecture, and one weekly meeting inquiry-based group work.

In the inquiry-based treatments, each student turned in each class meeting a written report on his/her investigation and solution of the problem(s) posed in that class period. This report is evaluated based upon the same rubric as the open-ended items on the pre/post-test (see Appendix 1 for a copy of the rubric). Students were aware of the rubric and received written feedback consistent with the rubric. In the lecture treatment, the instructor gave a traditional lecture on the upcoming material. All instructors operated from the same outline of topics for each lecture. The 14% (140 of 1000 points) of the final grade determined by the class meetings differed among the three treatments as follows:

GG: 5 points are earned for each of the two weekly reports on the group work;

LL: 5 points are earned for attendance at each class meeting;

GL: 5 points are earned for the one weekly report on the group work meeting, and 5 points are earned for attendance at the lecture meeting.

This research was carried out in Fall, 2010. Data gathered included (1) course grades and test scores, (2) pre-test and post-test of content knowledge based upon a test which incorporates three open-ended problems, evaluated on rubric dimensions of conceptual understanding, evidence of problem-solving, and adequacy of explanation (3) pre-test and post-test of content knowledge based upon a test consisting of 25 objective questions, and (4) student course evaluations using the online IDEA system (IDEA, 2010), and (5) RTOP observations of the instructors in each of the nine class sections (RTOP, 2010; Sawada, 2002).

A limitation of the studies by Mayer (2009, 2010) is that the pre/post-test consisted of only three or four open-ended problems which made a reliable evaluation of accuracy gains, if any, problematic. The pre/post-test in the study described herein consists of two parts:

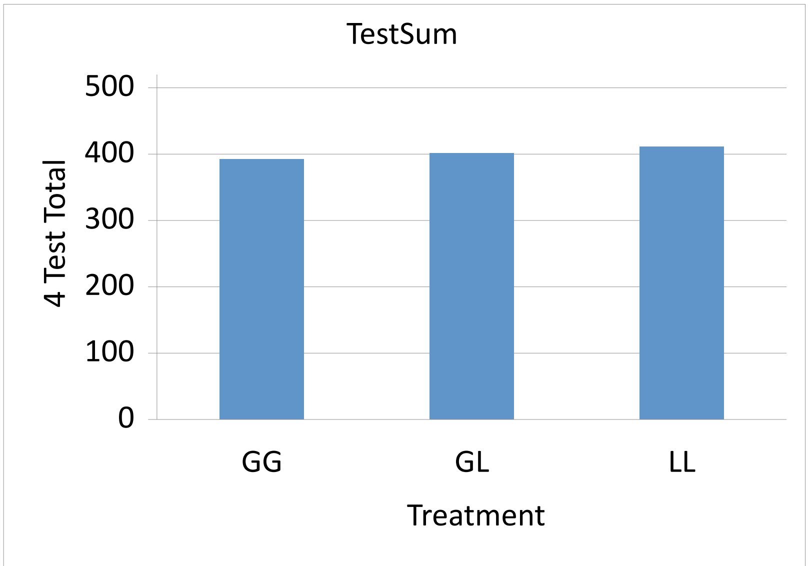
Part 1: three open-ended problems, evaluated by a rubric as described above, and

Part 2: 25 objective questions which had been validated for testing algebraic content knowledge in previous studies.

A battery of the previously validated (for content) objective questions was piloted in Summer 2010 on students in the same course, and item analysis was used to select the items for the pre/post-test in this study. As a result of the more careful test design, we expected that differential gains in accuracy between treatments, if present, would be more detectable than in the two earlier studies cited.

Results

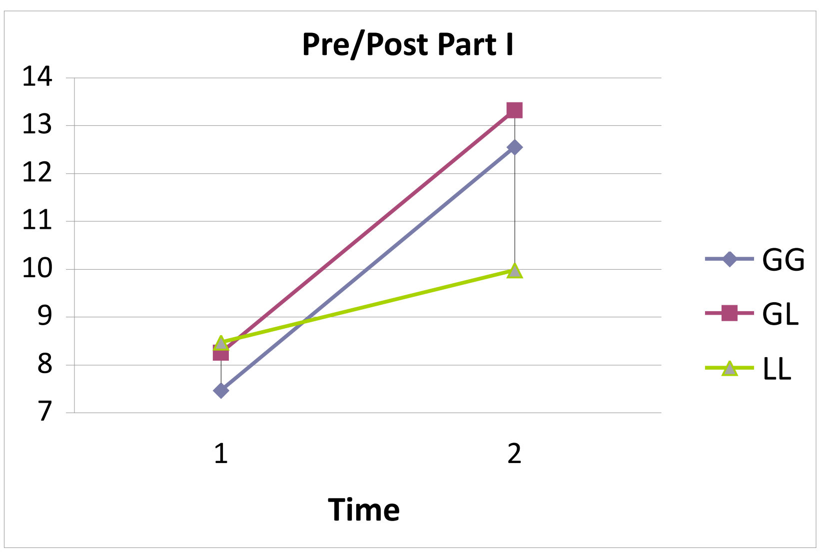
Students in all three treatments take the same five tests during the term. The tests are administered in the mathematics learning lab and are graded by computer. Tests are short answer rather than multiple choice. Each test is worth 130 points. Since the tests are the most significant determiner of student grades in the course, we used the sum of the first four of the five tests as our measure of the impact of the treatments on student grades. The maximum possible score is thus 520. The following graph shows the average test sum by treatment.



(N=315, GG=100, GL=106, LL=109.)

There was no significant difference ($p<0.05$) in the sum of the four test grades between any pair of treatments. Nor were there significant differences on any single test. The first four tests were used because some students who are satisfied with their accumulated course grade prior to test 5 elect not to put forth much effort on the fifth test. (A test sum of 400 points on the first four tests is well on the way to a "B" in the course, other grade factors being typical.)

Part 1 of the pre/post-test consisted of three open-ended problems graded by a four-part rubric. On each part of the rubric (see Appendix 1 for a copy of the rubric), a student could score up to two points, for a total of 8 points per question, and a total of 24 points on Part 1 of the test. The graph below shows the results by treatment on Part 1 of the pre/post-test.

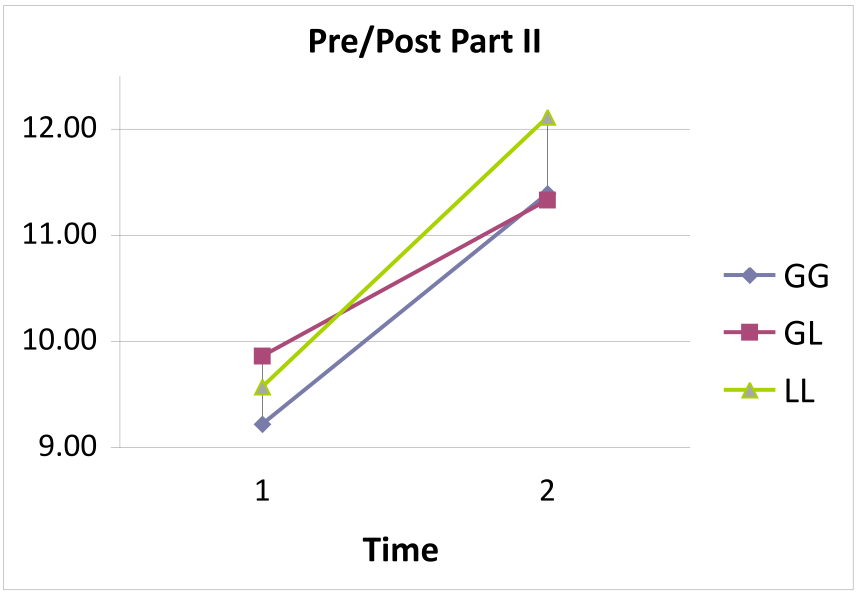


(N=272; GG=85; G =93; LL=94.)

One way ANOVA on the pre-test scores showed that there was no significant difference ($p<0.05$) between any pair of treatments. The GG and GL treatments differed significantly from the LL treatment. Repeated measures ANOVA showed that both the Time effect and the

Time*Treatment interaction effect were statistically significant ($p<0.05$). Wilkes Lambda for the Time effect was $\lambda=0.690$ and for the Time*Treatment interaction effect was $\lambda=0.921$. One way ANOVA on the post-test scores confirmed this finding: there was no significant difference between the GG and GL treatments, but there were significant differences between the GG and LL and between the GL and LL treatments.

Part 2 of the pre/post-test consisted of 25 objective questions, some yes/no, some multiple choice, and some always/sometimes/never. The expected value of the test (answering at random) was 10.38. The following graph shows the results by treatment on Part 2 of the pre/post-test.



(N=273; GG =88; GL =91; LL =94.)

One way ANOVA on the pre-test scores showed that there was no significant difference ($p<0.05$) between any pair of treatments. The Time effect was statistically significant ($p<0.05$), and Wilkes lambda for the Time effect was $\lambda=0.759$. There was no significant Time*Treatment interaction effect. One way ANOVA on the post-test scores confirmed this finding: there was no significant difference between any pair of treatments.

We also computed the effect size (difference of means divided by the pre-test standard deviation) for each of the treatments. The effect sizes ranged from medium (>0.40) to large (>0.80). However, in the absence of any statistically significant treatment effect, it is difficult to know how to interpret the results concerning effect size.

	Mean	*Mean*	*Standard Deviation*	*Standard Deviation*	*Effect Size*
Treatment	Pre	Post	Pre	Post	
GG	9.22	11.39	3.02	2.98	0.72
GL	9.86	11.33	3.44	3.38	0.43
LL	9.57	12.11	3.00	3.32	0.84

The IDEA student evaluations produce both numerical data that can be compared to the IDEA database of courses nationwide and within the discipline (mathematics), as well as written comments from students. Most students who complete the survey do not write comments. The course instructors and two additional readers were given student comments identified only by a code letter blindly corresponding to treatment and asked to identify "themes" in the comments. Following are the themes identified by the two additional readers (which are consistent with those identified by the instructors).

Reader 1: (General comment: I didn't count those who just said "good teacher" among the like. I was more looking for comments on the format.)

GG: Never taught/no help very frequent (13); other negative (unspecific) (7); only two positive.

GL: Like/learned a lot very frequent (14); some think online not helpful (3).

LL: Helped/understood better (6); lab good/class bad (4); tests should be closer to when material is learned (2).

Reader 2: (General comment: These are my observations concerning the major themes in the three sets of student comments.)

GG: The lack of lectures was harmful (the teacher did not teach).

GL: The instructor was the key to learning and made a great difference in the value received.

LL: The math learning lab is very helpful and was a key element in learning.

Student comments on the GG treatment appeared to be overwhelmingly negative in the view of all the readers, including the three course instructors. The comments on the GL treatment were more "middle of the road" and the comments on the LL treatment were mixed, but generally positive.

The table below represents the IDEA survey overall ratings of student evaluation of instruction. The Raw Average in each category is on a 5-point Likert scale, where 5 represents "strongly agree" and 1 "strongly disagree."

	Raw Average		Converted
GG	Excellent Teacher	Excellent Course	Score
Instructor 1	2.4	2.2	25
Instructor 2	4.2	4.0	52
Instructor 3	2.5	2.6	30

	Raw Average		Converted
GL	Excellent Teacher	Excellent Course	Score
Instructor 1	3.7	3.6	45
Instructor 2	4.3	4.0	52
Instructor 3	4.8	4.1	56

	Raw Average		Converted
LL	Excellent Teacher	Excellent Course	Score
Instructor 1	4.3	3.9	51
Instructor 2	4.3	4.0	50
Instructor 3	4.3	4.0	49

The form in which we receive the IDEA ratings does not permit aggregating student responses by treatment. The Converted Score (an adjusted average of the two raw scores) allows one compare student evaluations of instruction to the IDEA database of all courses: scores in the 45-55 range place the course/instructor evaluation in the middle 40% of all student ratings. Scores 37 or lower are in the lowest 10% of all student ratings.

Conclusions

We summarize our conclusions based upon the above results as follows.

- The inclusion of group work class meetings in lieu of lecture does not appear to affect adversely student success as measured by grades.

- Inquiry-based group work does have a positive effect on problem-solving and communications abilities as measured by the rubric score (Part 1) of the pre/post-test.
- Inquiry-based group work does not appear to affect accuracy as measured by the objective part (Part 2) of the pre/post-test.
- Two group work sessions do not appear to be significantly better than one per week, as indicated by the statistical indistinguishabilty of the GG and GL treatments.
- Student satisfaction is moderate with both the LL and GL treatments, but students are unsatisfied with the GG treatment.

Implications

This research will inform our teaching of elementary algebra. We do not know if the results of this study are generalizable to other mathematics courses with differing content. However, we expect to extend this study in subsequent years to credit courses such as intermediate algebra, college algebra, and pre-calculus algebra and trigonometry (Oerhtman, 2008). The aforementioned courses all are computer-assisted. This further study will allow us to examine the cumulative effect on students of multiple classes incorporating inquiry-based sessions in a computer-assisted context.

All treatments had statistically indistinguishable test scores and course final grades. Any valued-added from including inquiry-based sessions is not being captured in course grades. This is not surprising given that 86% of the course grade is determined by the computer-assisted component. The rubric-scored Part 1 of the pre/post-test appears to capture at least part of the value-added: students are better able to express their thinking about mathematics problems. However, it is depressing that the objective Part 2 of the pre/post-test does not show any difference among treatments. It is possible that cumulative exposure to learning constructively will have a detectable objective effect. We are also continuing to study the pre/post-test itself, the beginnings of which are presented in Appendix 2.

Many of our students in elementary algebra did not appear to find value in the inquiry-based components. But they did, as witnessed by the reactions to the GL treatment, tolerate inquiry-based sessions when blended with traditional instruction. As instructors, we value the deeper thinking that can arise in learning constructively. We provide an illustrative story of this in Appendix 3. While we do not discount student satisfaction, as teachers, we ask ourselves if pedagogical decisions ought to be based on whether or not students perceive (in the short term) any value in the inquiry-based components of the learning process? If blended treatments are more accepted by students then this perhaps should be our direction in the short term.

References

Blais, D. M. (1988), Constructivism -- a Theoretical Revolution for Algebra. Mathematics Teacher, Nov. 1988: 49--56.

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

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

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

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

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

Mayer, J.C., et al. (2009), Incorporating Inquiry-Based Class Sessions with Computer Assisted Instruction. Electronic Proceedings of 2009 SIGMAA Conference on Research in Undergraduate Mathematics Education, Raleigh, NC, February, 2009.

Mayer, J.C., et al. (2010), Blending Inquiry-Based and Computer Assisted Instruction in a Basic Algebra Course: a Quasi-Experimental Study, Electronic Proceedings of Conference on Research in Undergraduate Mathematics Education, Raleigh, NC, February, 2010.

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

Pyrczak, F. (1973). Validity of the discrimination index. Journal of Educational Measurement 10(3): 227-231.

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

Sawada, D., Piburn, M., Judson, E., Turley, J., Falconer, K., Benford, R. & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. School Science and Mathematics, 102(6), 245-253.

Wiggins, G. (1989), The Futility of Trying to Teach Everything of Importance. Educational Leadership, Nov. 1989: 71-78.

Appendix 1

MA 098/110 Scoring Guide

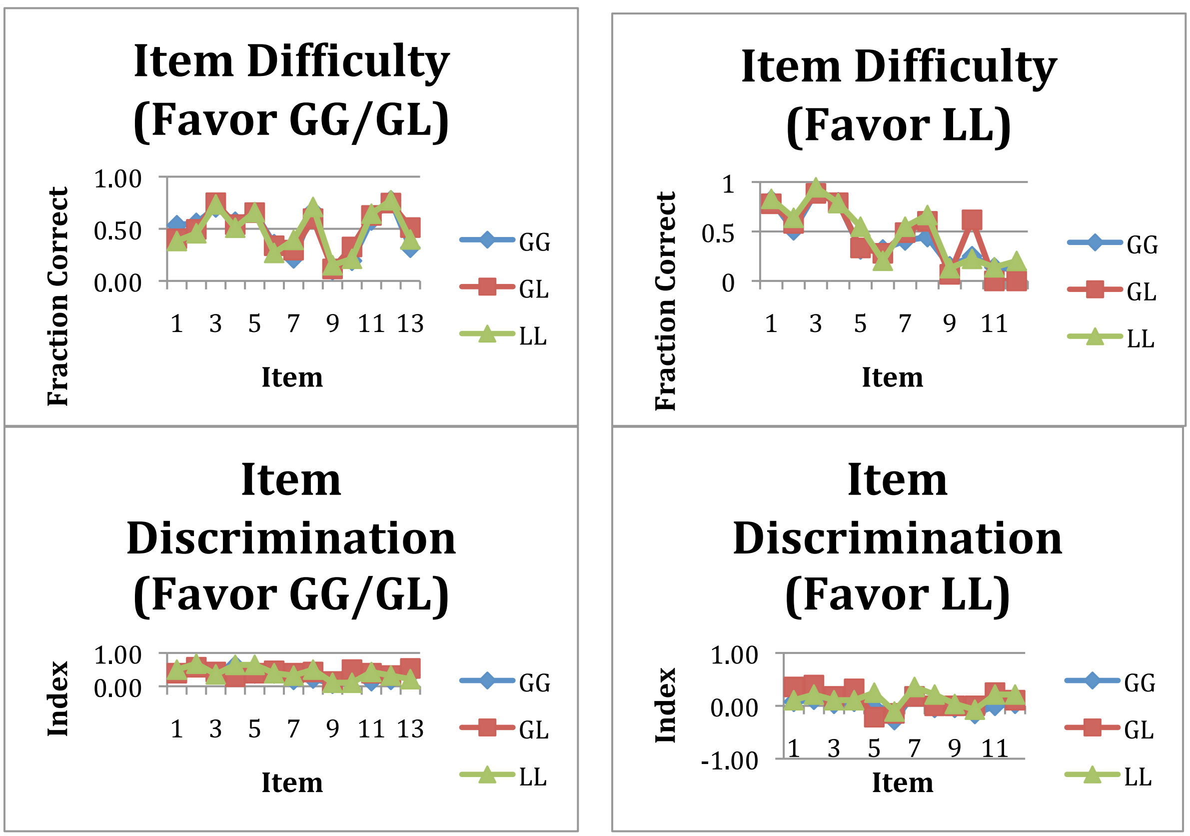
	Conceptual Understanding: *Interpreting the concepts of the task and translating them into mathematics*	**Evidence Of Problem Solving:** *Choosing strategies that can work, and then carrying out the strategies chosen.*	**Explanation:** *Using pictures, symbols, and/or vocabulary to convey the path to the identified solution*	**Accuracy:** *Providing a complete and accurate solution appropriate for the given problem*
2	The translation of the task into adequate mathematical concepts using relevant information is completed.	Pictures, models, diagrams, symbols, and/or words used to solve the task are complete.	Explanation is clear and complete.	Solution is correct and complete with no errors.
1	The translation of the major concepts of the task is partially completed and/or partially displayed.	Pictures, models, diagrams, symbols, and/or words used to solve the task may be only partially useful and/or partially recorded.	The explanation is partially complete and/or partially developed with gaps that have to be inferred.	Solution is appropriate and demonstrates understanding, but is either not quite complete or contains minor errors.
0	Does not achieve minimal requirements for 1 point.	Does not achieve minimal requirements for 1 point.	Does not achieve minimal requirements for 1 point.	Does not achieve minimal requirements for 1 point.

Loosely adapted from the Oregon Department of Education's 1995-2003 statewide assessments (originally a 5-point rubric).

Appendix 2. Preliminary Analysis of Part 2 of Pre/Post-Test.

Because our expectation of improved performance on the objective part (Part 2) of the pre/post-test by those receiving more inquiry-based instruction was not realized, we have begun a more detailed analysis of Part 2. Here we present some preliminary results.

To begin, Instructors/TAs were asked to rate on a scale of 1 (lowest) to 5 (highest) the extent to which students received instruction that would help them in responding to each pre/post-test question, with each treatment being considered separately. Based upon the responses, we tentatively identified 13 questions on Part 2 as possibly favorable to students in the GG or GL treatments and 12 as indifferent or possibly favoring the LL treatment. We then performed item difficulty analysis and a simple item discrimination analysis (Pyrczak, 1973) on each so-identified group of questions separately, with respect to each treatment. The following graphs illustrate the results.

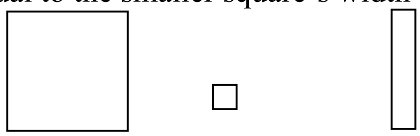


No clear pattern emerges and further study is required.

Appendix 3: An Example of a Group Learning Situation

The group learning experience is vastly different than the traditional class lecture experience. The following example of a group learning situation in one of the group work sessions illustrates this difference.

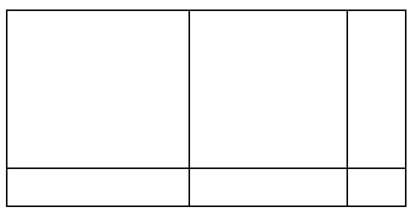
Students (working in groups of four) were given a set of algebra tiles at the beginning of a group work session. There are three different-sized pieces in the algebra tile sets. There are large square pieces, small square pieces, and (non-square) rectangular pieces. The rectangular piece has width equal to the smaller square's width and length equal to the larger square's width.



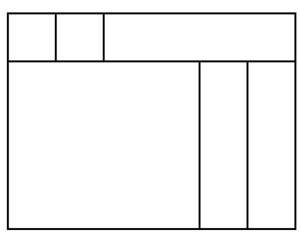
Students were given a warm-up problem using a set of these tiles. Students were asked to view the three tile pieces as a 10 x 10 square, a 1 x 10 square, and a 1x 1 square. Then students were asked to show a rectangular area model of 264 using the tiles.

After the warm-up problem, students were asked to construct the following figures: an x-by-x square, a 1-by-x square, and a 1-by-1 square and label the area of each of these figures. Next, students were directed to represent a trinomial such as as a rectangular area model using the tiles. An overwhelming majority of students use the larger square from the algebra tiles as the x-by-x square and the smaller square as the 1-by-1 square.

Then they create a rectangular area model similar to the one below.



In contrast to the rest of the students, one student in the class used the smaller square as the x-by-x square and the larger square as the 1-by-1 square. This student's rectangular area model is shown below.



The student and his group struggled with the idea of why there were two different models representing the same algebraic expression. After about twenty-five minutes and some questions from the instructor, the student who chose the "unusual" depiction of realized that he could put certain restrictions on the value of the variable (x >1 or x <1) to explain the two different representations. This student and his group initially expressed concern that his model was different and assumed that only one model could be correct.

This type of experience is unlikely to occur in a lecture setting. Students do not have the opportunity to explore different scenarios and models for a problem in a lecture. The example above shows how students can arrive at a deeper understanding of a concept through encountering unexpected outcomes and working through the reasoning behind these unexpected outcomes.