Enacting Reasoning-and-Proving in Secondary Mathematics Classrooms through Tasks

by

Michelle S. Switala

B.A. in Physics, Franklin & Marshall College, 1991

M. S. in Physics, University of Arizona, 1994

Submitted to the Graduate Faculty of

The School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Education

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

The School of Education

This dissertation was presented

by

Michelle S. Switala

It was defended on

July 29, 3013

and approved by

Dr. Ellen Ansell, Associate Professor, Instruction and Learning

Dr. Melissa Boston, Associate Professor, Duquesne University

Dr. James Greeno, Emeritus Professor, Stanford University

Dissertation Advisor: Dr. Margaret S. Smith, Professor, Instruction and Learning

Copyright © by Michelle S. Switala

2013

Enacting Reasoning-and-Proving in Secondary Mathematics Classrooms through Tasks

Michelle S. Switala, EdD

University of Pittsburgh, 2013

Proof is the mathematical way of convincing oneself and others of the truth of a claim for all cases in the domain under consideration. As such, reasoning-and-proving is a crucial, formative practice for all students in kindergarten through twelfth grade, which is reflected in the Common Core State Standards in Mathematics. However, students and teachers exhibit many difficulties employing, writing, and understanding reasoning-and-proving. In particular, teachers are challenged by their knowledge base, insufficient resources, and unsupportive pedagogy.

The Cases of Reasoning and Proving (CORP) materials were designed to offer teachers opportunities to engage in reasoning-and-proving tasks, discuss samples of authentic practice, examine research-based frameworks, and develop criteria for evaluating reasoning-and-proving products based on the core elements of proof. A six-week graduate level course was taught with the CORP materials with the goal of developing teachers' understanding of what constitutes reasoning-and-proving, how secondary students benefit from reasoning-and-proving, and how they can support the development of students' capacities to reason-and-prove. Research was conducted on four participants of the course during either their first or second year of teaching. The purpose of the research was to study the extent to which the participants selected, implemented, and evaluated students' work on reasoning-and-proving tasks. The participants' abilities were examined through an analysis of answers to interview questions, tasks used in class, and samples of student work, and scoring criteria. The results suggest that: 1.) participants

were able to overcome some of the limitations of their insufficient resource by modifying and creating some reasoning-and-proving exercises; 2.) participants were able to maintain the level of cognitive demand of proof tasks during implementation; and 3) participants included some if not all of the core elements of proof in their definition of proof and in their evaluation criteria for student products of reasoning-and-proving products.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Seven years is a long journey, and there were times when I never thought I would finish. I am deeply indebted to so many people that helped me along the way. First and foremost, I would like to thank my parents, my siblings and their spouses in education (Dr. Kristen, David, Bill, Jen, Kevin, soon-to-be Dr. Katie, John, Maria, Brian, Sarah, Michael, and Gina), Aunt Barbie, Uncle Dan, Aunt Krista, and my nieces and nephews. Dr. Richard McNabb, thank you for your help and support over the last 20 years—I finally caught up! I would also like to thank my fellow graduate students: Leah, Milan, Justin, Scott, Steve, and Josh, but especially Sam for his help and humor in the office, Marcella for grounding and inspiring me, and Mary Ellen, for her intelligence, drive, and friendship. Mary Ellen, I look forward to discussing education and teaching with you for many years to come.

I would also like to thank my committee for their attention and good advice during this process, especially my advisor, Dr. Peg Smith. I appreciate your time, funding, and invitation to be part of the CORP project. Dr. Melissa Boston, your hard work on the IQA Mathematics Toolkit helped make my study possible, and your kindness and humanity helped me complete my dissertation. Dr. Ellen Ansell, thank you for the close read of my document. Dr. Jim Greeno, thank you for considering me worthy of conversation about education and research.

There are many people from the public school community that inspired and encouraged me before and during this process: Dr. Laura Davis, Dr. Ron Meisberger, Drs. Bob and Jeanne

Johnson, Dr. Bille Rondinelli, Dr. Mike Loughead, Dr. Ed Sarver, and Dr. George Szymanski. The conversations we have had about how to help students were formative, interesting, and real. I would not have embarked on this process without you, and I hope to be half the educator that you are. Thank you to my fellow teachers, especially Eric, Kurt, and the NSTOY-PA Board members who offered themselves as sounding boards and were great sources of encouragement. I would also like to thank my Pine-Richland parents who made me meals, sent me good wishes, and were excited about my progress.

Finally, I dedicate this study to my students. You are, and always will be, worth the effort.

1.0 STATEMENT OF PROBLEM

1.1 INTRODUCTION

Every discipline has a method for verifying conclusions drawn from observations, data, and analysis. The method mathematicians use to draw valid conclusions from conjectures is *proof* (Hanna & Jahnke, 1993). Proof is the mathematical way of convincing oneself and others that an idea is absolutely true for all cases in the domain under consideration. According to the three major sets of standards for mathematics education (NGA Center & CCSSM, 2010; NCTM, 1989, 2000), proof is a crucial, formative practice for all students in kindergarten through twelfth grade. The *Principles and Standards for School Mathematics* (NCTM, 2000) states that mathematics education should help students see that claims must be validated by the community at hand, experimentation and conjectures can lead to discovery, and arguments must be clearly communicated. The more recent Common Core State Standards for Mathematics (NGA Center & CCSSO, 2010) incorporates proof slightly differently in its *Standards for Mathematical Practice:* mathematically proficient students can "reason abstractly and quantitatively; construct viable arguments and critique the reasoning of others; and look for and make use of structure" (p. 6-8). Both sets of standards advocate incorporating reasoning and proving across the curriculum. Since students learn at different rates (Senk, 1985) and the curricular development of proof does not need to be linear (Hoyles, 1997; Knuth, 2002c), the habit of reasoning mathematically should be developed consistently over time in many contexts (NCTM, 2000; Senk, 1985).

1.2 STUDENT DIFFICULTIES WITH PROOF

Student understanding of mathematical proof can be described in terms of the types of connections that students have formed between definitions, theorems, and procedures (Hiebert et al., 1997). Research has shown, however, that secondary and post-secondary students of all ability levels and courses have a difficult time coordinating definitions, theorems, and procedures in order to make connections and construct valid arguments (Bell, 1976; Coe & Ruthven, 1994; Galbraith, 1981; Healy & Hoyles, 2000; Ko & Knuth, 2009; Senk, 1985). Proof-making is a very cognitively demanding task. Yackel and Hanna (2003) explained these demands:

> By its very nature, mathematical proof is highly sophisticated and seems to be much more challenging intellectually than many other parts of the school mathematics curriculum. To a large extent this is so because the kind of reasoning required in mathematical proof is very different from that required in everyday life. Reasoning in everyday life does not require the rigor of mathematical proof nor the careful attention to process demanded by the mathematical proof, which seeks, for example, to make clear distinctions among assumptions, theorems, and rules of inference. (p. 231).

In addition to making distinctions between and making connections among assumptions, definitions, theorems, etc., students must also be able to access relevant background knowledge, maintain the integrity of the concept of proof, and craft a convincing argument (Harel & Sowder, 2007; Healy & Hoyles, 2000).

Students also struggle with deductive reasoning skills. Consider the two branches of reasoning: induction and deduction. In inductive reasoning, a hypothesis is made, phenomena are observed, and a conclusion is stated which applies to all cases that are defined for the observed phenomena. Since every instance of a case cannot be actually observed, the conclusion is based on a discerned pattern from the phenomena. As such, inductive reasoning may be a step in the proving process but does not constitute a proof itself. Deductive reasoning, however, is a logical process in which the conclusion does *not* contain more information than the collection of premises stated in the beginning of the process; an argument based on deductive logic may constitute a proof. Even with this knowledge, however, students tend to favor empirical arguments based on inductive reasoning (Chazan, 1993; Coe & Ruthven, 1994; Healy & Hoyles, 2000). This could be because outside of mathematics, "proof" is synonymous with evidence (Harel & Sowder, 2007; Healy & Hoyles, 2000; Hersh, 1993; Yackel & Hanna, 2003), or students' limited mathematical knowledge (e.g., distinguishing theorems, accessing relevant knowledge, precisely defining terms) makes writing deductive proofs a generally unsuccessful endeavor (Bell, 1976; Chazan, 1993; Chazan & Lueke, 2009; Coe & Ruthven, 1994; Galbraith, 1981; Healy & Hoyles, 2000; Moore, 1994; Senk, 1985). Deductive reasoning requires strict attention to precise definitions and stated postulates and theorems. Students' ability to successfully launch into a deductive proof depends on their intuition (Chazan, 1993; Yackel & Cobb, 1996), investigation techniques (Coe & Ruthven, 1994; Senk, 1985), and ability to chain logical inferences (Bieda, 2010; Galbraith, 1981; Hanna & Jahnke, 1993; Harel & Sowder, 2007; Senk, 1985). Another possibility is that the goal of a practical person is to be efficient (Balacheff, 1991; Coe & Ruthven, 1994; Schoenfeld, 1985) and achieving rigor and systematization are generally not efficient processes. Students do appear to be better at choosing correct proofs than constructing their own (Healy & Hoyles, 2000), perhaps because generating a sequence of steps demands different cognitive skills than understanding someone else's proof (Moore, 1994).

Choosing a reasonable format for a proof further complicates the process for students. Some students forgo a format that clarifies and communicates an argument for them in favor of a formal, abstract argument that they believe the teacher desires (Chazan, 1993; Healy & Hoyles, 2000; Hoyles, 1997; Knuth, 2002b; Küchemann & Hoyles, 2001). Unfortunately, when form is seen as paramount, students can create or accept arguments that are illogical or fail to convince or explain an argument (Balacheff, 1991).

Students also frequently misunderstand the role of proof in secondary mathematics. Since proof is the mathematical way of convincing oneself and others that an idea or concept is absolutely true for all cases in the domain under consideration, students must see the need for *convincing* in order to see the need for *proof*. Some students assume that every theorem encountered in a textbook has already been proven and is thus to be blindly accepted; therefore, there is no need for the students themselves to prove the theorem (Chazan & Lueke, 2009; Hanna & Jahnke, 1993; Tinto, 1990). There is no need for a proof if everyone is convinced of the truth of a statement. According to some students, only investigations require proof, and that is only because teachers expect them (Harel & Sowder, 2007; Hoyles, 1997; Mariotti, 2000). Healy and Hoyles (2000) found that 50% of surveyed students thought the purpose of proof was to obtain truth, 35% thought the purpose was to explain a concept, 1% thought proof was about discovery or systematizing, and a full 28% thought proof had no purpose or could not think of one. Even when students accept a proof task and are able to construct a valid argument, many students fail to see the generality of the proof (Chazan, 1993; Chazan & Lueke, 2009; Galbraith, 1981;

Martin, McCrone, Bower, & Dindyal, 2005) or hesitate in accepting the proof because it was based on assumptions (Chazan, 1993), which loops the students back to misunderstanding the role of proof in mathematics. In summary, students have difficulty with rigor, process, and purpose of proofs, coordinating assumptions, definitions, and theorems, accessing relevant background knowledge, crafting a convincing argument in a reasonable format, and using deductive reasoning.

1.3 TEACHER DIFFICULTIES WITH PROOF

Teachers have a responsibility to help students with all of these difficulties, but research has shown that many teachers exhibit some of the same difficulties as their students. Some teachers have limited knowledge of proof (Knuth, 2002b; Stylianides, A. J. & Stylianides, G. J., 2009), their resources are insufficient (Johnson, Thompson, & Senk, 2010; Senk, 1985; Stylianides, G. J., 2009), and their pedagogy does not generally support students' development of reasoning-and-proving (Bieda, 2010; Chazan & Lueke, 2009; Ellis, 2011; Lampert, 1990). Knuth (2002b) found that like students, some teachers misunderstood the *generality* of a completed proof. Unlike mathematicians, some teachers thought that a proof merely verified the truth of a theorem rather than explained *why* the theorem was true. In addition, some teachers exhibited weak deductive reasoning skills. When presented with a list of arguments, the teachers in Knuth's study rated 33% of non-proofs as proofs, some by focusing on the algebraic manipulations rather than on the logical validity of the argument. Unfortunately, curricula materials offer little help (Johnson, Thompson, & Senk, 2010; Senk, 1985; Stylainides, G. J., 2007). Johnson, Thompson, and Senk (2010) found in their study of selected chapters of algebra and precalculus textbooks that only 5.5% of textbook problems required reasoning and 46% of proof-related reasoning problems were based justifications on a single case. It is worthwhile to note that this study looked at textbooks in wide circulation a full ten years after NCTM's *Principles and Standards for School Mathematics*.

In many classrooms the teacher—expert or not—is the sole validation authority (Chazan & Lueke, 2009). When an *authoritative proof scheme* is created in a classroom (Harel & Rabin, 2010), students expect to be told information rather than work to construct information for themselves, resulting in helpless students who are reluctant to question assertions (Harel & Sowder, 1998). In such a system, "*Doing* mathematics means following the rules laid down by the teacher; *knowing* mathematics means remembering and applying the correct rule when the teacher asks a question, and mathematical *truth* is determined when the answer is ratified by the teacher" (Lampert, 1990, p. 31). Bell (1976) wrote that proof is a public activity, and "pupils will not use formal proof with appreciation of its purpose until they are aware of the public status of knowledge and the value of public verification" (p. 25).

Beliefs, knowledge, instruction, and classroom management all impact a teacher's ability to provide reasoning-and-proving opportunities to students. Research suggests that what students learn depends greatly on the beliefs and actions of the teacher (Harel & Sowder, 2007; Yackel & Cobb, 1996). If a teacher believes that all students can learn, then the teacher structures classroom interactions to include all students. If a teacher believes that only the best mathematical students can learn deductive logic, then the teacher grapples with the tension of helping enculturate students into the field of mathematics while lowering expectations to accommodate perceived abilities (Martin et al., 2005; Bell, 1976). This in turn conveys expectations to students for how knowledge is communicated (Herbst, 2002). Students experience mathematics through the tasks, methods, and instructional strategies that the teacher chooses (Harel & Sowder, 2007; Martin et al., 2005). For example, when teachers communicate that an efficient way to solve problems or construct proofs is to remember the work of others or reduce the problem into computational rather than logical terms (called *smoothing out the curriculum* by Doyle in 1988), opportunities for students to develop autonomous learning capabilities are limited and should be avoided. Clearly, teachers need to make careful pedagogical decisions in order to promote the development of reasoning-and-proving skills (Stylianides, G. J. & Stylianides, A. J., 2009).

1.4 PROFESSIONAL DEVELOPMENT

Since an emphasis on developing connections and communicating discoveries is vastly different than how most teachers were trained (Ball, 1988b), professional development is essential in helping teachers develop their subject-matter knowledge so that they can support their students' development of reasoning-and-proving understanding and skills. Professional development with a focus on enacting reasoning-and-proving should help teachers overcome difficulties with their knowledge of proof, insufficient resources, and unsupportive pedagogy. The materials developed by the NSF-funded Cases of Reasoning-and-Proving (CORP) project were designed to provide teachers with exactly these types of opportunities. The CORP work utilizes Ball and Cohen's (1999) work on grounding teachers' professional development in tasks and problems of practice. Specifically, the CORP materials provide teachers with research-based frameworks on tasks and reasoning-and-proving, opportunities to create criteria for evaluating proof, samples of student work to analyze, rich narrative cases to discuss, and proof tasks on which to work.

Through the use of these types of materials, teachers develop their knowledge of mathematics for teaching (Ball, Thames, & Phelps, 2008).

The participants in the study that will be described herein are teachers who completed a reasoning-and-proving course based on the CORP materials while they were preservice teachers. The teachers were followed into their classrooms to see the extent to which they engaged their students in reasoning-and-proving through the tasks the teachers selected and implemented and the extent to which the teachers evaluated their students' reasoning-and-proving products with the core elements of proof.

Hiebert et al. (1997) wrote about five dimensions that help define classrooms in terms of the learning opportunities for students: nature of tasks, role of teacher, social culture of the classroom, availability of mathematical tools, and participation of all. The reason this study focused on tasks is that tasks influence students' opportunities to learn more than any other dimension (Lappan & Briars, 1995). "A *mathematical task* is defined as a classroom activity, the purpose of which is to focus the students' attention on a particular mathematical idea" (Stein, Grover, & Henningsen, 1996, p. 460). Mathematical tasks have the potential to clarify students' concept of proof and help them engage in reasoning-and-proving. When selecting tasks, teachers should consider the problematic aspects of the task, the developmental state of their students, and the mathematical goals (Hiebert et al., 1997). The tasks must be interesting enough for students to want to solve and discuss the tasks (Hiebert et al, 1997), and sets of tasks should allow students to build their knowledge through different points of view (Simon, 1995).

Teachers must strike a balance between maintaining the high-cognitive demand of good tasks and providing necessary help. Stein et al. (1996) identified four levels of cognitive demands associated with a task: memorization, use of procedures and algorithms (with or without understanding), and complex thinking and reasoning strategies. The researchers also found three distinct phases in enacting tasks in classrooms: set-up by the teacher, implementation by the teacher and students, and resulting student learning. The set-up was influenced by teacher goals, subject-matter knowledge, and knowledge of students. The implementation was influenced by classroom norms, task conditions, teacher instructional habits and dispositions, and student learning habits and dispositions. Rich learning experiences occurred for students when: (a) tasks were cognitively demanding and built on students' prior knowledge, (b) teachers provided scaffolding and an appropriate amount of time as well as modeled high-level performance, maintained pressure for explanations and meaning, and drew conceptual connections, and (c) students self-monitored. All of these activities help students coordinate competencies and see the value of proof.

When a teacher has a good grasp of the mathematical ideas involved in the proof, the teacher is in a better position to provide information such as mathematical conventions for recording ideas and alternate solution paths and to highlight important mathematical ideas in students' work (Hiebert et al., 1997). Teachers who provide this information are acting as a conduit between learners and the professional mathematics community; they are not smoothing out the curriculum. G. J. Stylianides (2008) created an analytical framework for the learning activities involved in reasoning-and-proving which can help teachers plan a learning trajectory for their students. Stylianides described a hierarchy of activities which start with reasoning (identify a pattern then make a conjecture) and end with providing support for mathematical claims (non-proof arguments or proof-arguments). For instance, an Algebra 2 teacher may want to incorporate reasoning-and-proving into her lesson on transformations of functions. The task selected from the textbook might merely ask students to sketch $|x+2|$. The teacher, with

Stylianides' framework in mind, might modify the task to "what can you say about functions of the type $f(x + a)$? How do you know?" The student could choose a function, such as the absolute value function, explore several instances of the type |x+a|, make a conjecture, then seek to prove that the conjecture works for any type of function. The student should be left with some knowledge of mathematical practice (such as proof) beyond the solution of the task, an idea called "residue" (Davis, 1992). Residue could be an insight into the structure of mathematics (Cobb, Wood, Yackel, Nicholls, Wheatley, Trigatti, & Perlwitz, 1991; Fuson & Briars, 1990; Hiebert & Wearne, 1993) or strategies or methods for crafting proofs (Fennema, Franke, Carpenter, & Carey, 1993; Hiebert and Wearne, 1993; Wearne & Hiebert, 1989).

There is no question that reasoning-and-proving is an essential practice of mathematics. There is also a large body of empirical research that shows many areas of weakness among students and teachers when it comes to proof. Unless teachers receive training or professional development on how to better teach reasoning-and-proving by implementing cognitively-demanding sets of proof tasks, offering appropriate and quality help, and creating a classroom culture in which everyone shares authority, students will probably remain impoverished in their understanding of proof and reasoning.

1.5 PURPOSE OF THIS STUDY

The purpose of this study was to investigate the extent to which teachers who participated in professional development related to reasoning-and-proving are able to select, modify, implement, and evaluate the student products of reasoning-and-proving tasks. The teachers' abilities will be examined through answers to interview questions, tasks used in class, and student work packets with accompanying rubrics.

1.6 RESEARCH QUESTIONS

This study analyzed the opportunities teachers gave their students to engage in reasoning-and-proving, after the teacher received professional development targeting reasoning-and-proving. The targeted professional development involved learning about the importance of reasoning-and-proving in mathematics, using frameworks and criteria to examine and evaluate student work, and implementing reasoning-and-proving tasks. This study sought to learn how the teachers used this knowledge in the context of their classrooms. In particular, this study examined the following questions:

1. **To what extent did teachers select reasoning-and-proving learning opportunities in the form of tasks?**
    a. To what extent did the textbook include tasks that had the potential to engage students in reasoning-and-proving?
    b. To what extent did the teacher select tasks for instruction that had the potential to engage students in reasoning-and-proving?
    c. To what extent did the teacher modify tasks to affect the tasks' potential to engage students in reasoning-and-proving?
    d. What were the sources of the tasks that teachers selected for instruction?
2. **To what extent were teachers able to maintain the level of cognitive demand of the reasoning-and-proving task during implementation?**

3. **To what extent were teachers able to accurately evaluate their students' reasoning-and-proving products?**
    a. To what extent did teachers' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?
    b. To what extent did teachers apply the core elements of proof in evaluating their students' reasoning-and-proving products?
    c. In what ways did teachers communicate expectations regarding what was required to produce a proof to students?

1.7 SIGNIFICANCE

This study assumed that students develop reasoning-and-proving skills and understanding through the learning opportunities provided to them by their classroom teachers. This study hypothesized that teachers were better able to engage their students in reasoning-and-proving after participating in targeted professional development that used problems of practice. If teachers can improve their ability to recognize, select, modify, and implement cognitively demanding reasoning-and-proving tasks and evaluate the products of those tasks by focusing on the core elements of proof, then their students will experience quality learning opportunities. The results of this study contribute to the body of knowledge from research on teacher education and reasoning-and-proof. It has the potential to suggest how teachers internalize professional development and use those skills and knowledge in the context of their classroom.

1.8 LIMITATIONS

This study has several limitations. The four participants in the study constituted a small sample of convenience; the participants were not intentionally different. The participants who volunteered for this study were from a pool of former preservice teachers who participated in the reasoning-and-proving course at a tier-one research university. All four students had similar, but not exactly the same, course experience. Two of the participants took the course in the summer of 2011 with one instructor, and the other two participants had the course in the summer of 2012 with a different instructor (the primary researcher in the current study). Both instructors of the course were members of the curriculum development team who also modified the course slightly between the two courses, although the portion of the course involved with the current study was largely stable between the two years.

Because the primary researcher in the current study was an instructor of the course, the primary researcher knew two of the participants and maintained a professional relationship with them as they began their teaching careers. Despite that relationship, the primary researcher established protocols for interviews and collecting data that were applied across all parts of the data collection and analysis.

The data was collected remotely over the course of two months and was not sampled across the year. The data collected could under- or over-sample the participants' activities involving reasoning-and-proving. Because data was collected remotely and no observations were made of the participants with their students in their schools, some aspects of the participants' practice—such as lesson plans or classroom discourse—could not be reliably studied due to the high rate of inference that would be required by evaluating these through selected tasks, evaluation rubrics, and student work.

1.9 OVERVIEW

This document is organized into five chapters. This chapter argued the need for targeted professional development to help teachers overcome their difficulties with reasoning-and-proof and offer their students learning opportunities in the form of tasks to develop their reasoning-and-proving skills and understanding. The second chapter reviews research conducted on the nature of proof, obstacles teachers face with respect to reasoning-and-proving, professional development, and tasks. Chapter Three describes the data sources, coding, and analysis of the methodology of this study. Chapter Four describes the results of the analysis, and Chapter 5 discusses the findings, conclusions, and outlines suggestions for future research.

2.0 LITERATURE REVIEW

As proof is a critical aspect of mathematics, it should not be divorced from secondary mathematics education (Schoenfeld, 1994; Wu, 1996). Recent standards documents (NGA Center & CCSSO, 2010; NCTM, 2000) have acknowledged proof's place in the study of mathematics and have provided some guidance, but teachers still struggle with offering their students opportunities to develop a conceptual understanding of proof and to develop skills for constructing proof (Knuth, 2002b). Teaching proof puts high demands on teachers (Chazan, 1993; Herbst, 2002), not only because there is no consensus in the field of mathematics about the definition or role of proof in secondary mathematics education (Weber, 2008), but also because many teachers[1] struggle with constructing proof themselves (Knuth, 2002b) and students find the study of proof difficult (Bell, 1976; Chazan, 1993; Healy & Hoyles, 2000; Senk, 1985). The study described herein focused on aspects of proof that makes the teaching of proof challenging. As such, student difficulties with proof will only be discussed in relation to the work that teachers need to do to provide learning opportunities for their students to meet some of those challenges. In this chapter, the role and nature of proof, an analytical framework for reasoning-

[1] Throughout this chapter, the term "teachers" refers to practicing secondary teachers. The term "students" refers to middle school, high school, or undergraduate (non-specific, math major, preservice elementary or secondary mathematics) learner of mathematics. Longer labels (e.g., undergraduate student who is a preservice elementary mathematics teacher) will only be used if the longer label contributes something specific to the sentence. For a list of empirical studies and the studies' subjects as labeled by their authors, see Appendices A and B.

and-proving, aspects of proof that challenge teachers, professional development, and examining practice are discussed.

2.1 THE NATURE OF PROOF

2.1.1 The Nature of Mathematics and Definitions of Proof

The nature of mathematics has evolved over time. For centuries, the deductive reasoning and axiomatic system of the ancient Greeks reigned and was considered sufficient (Harel & Sowder, 1998; Hanna & Jahnke, 1993). During the sixteenth through eighteenth centuries, however, mathematicians relaxed their rigorous standards for validating theorems until the mathematicians developed the techniques of new fields such as algebra, analytical geometry, and calculus. This allowed them to resolve issues with all classes of numbers (e.g., irrational numbers) and some Greek paradoxes (e.g., Euclid's fifth postulate) (Boyer, 1991; Hanna & Jahnke, 1993; Harel & Sowder, 2007). In the nineteenth century, mathematicians had made enough progress to allow the return of rigorous axiomatic systems. Soon afterwards, the field experienced a crisis in philosophy regarding the nature of mathematics (Boyer, 1991). David Hilbert—a formalist—"considered mathematical objects to be symbolic entities which owe their existence only to the fact that they satisfy the rules by which they are axiomatically linked" (Hanna & Jahnke, 1993, p. 425). In such a field of objects, some derived theorems were incomprehensible until applications could be found (Davis & Hersh, 1981; Hanna & Jahnke, 1993). In contrast, Henri Poincare—who approached mathematics from a practical perspective—preferred to think of mathematics as being led by intuition and exploration (Schoenfeld, 1986). Currently, most mathematicians describe the field of mathematics as embracing a deductive system with clearly stated assumptions, using variables free from real or imagined referents, and with constructs that are open to interpretation (Harel & Sowder, 2007).

Since the role of proof evolved in conjunction with the philosophy of mathematics without the fusion of theoretical and practical mathematics (Boyer, 1991; Davis & Hersh, 1981; Hanna & Jahnke, 1993; Harel & Sowder, 2007; Schoenfeld, 1986), the field accepts three different perspectives on the definition of proof (Weber, 2008). The first (and most rigorous) definition described proof as a formal structure whose test of validity includes checking against well-defined, stated rules and conventions (Griffiths, 2000). A second definition was provided by Davis and Hersh (1981), who described a proof as a mathematician-convincing argument, although Davis and Hersh acknowledged that the reputation of the author may influence the acceptance of the proof. Finally, others (Balacheff, 1987; Manin & Zilber, 2010; Thurston, 1994) described proof as an argument to be negotiated socially. The last definition seems the most reasonably within the grasp of students who do not usually have the knowledge of or access to professional mathematicians (Balacheff, 2010). Regardless, most researchers agree that a mathematical proof at any level must be concerned with the truth of a statement, convey insight into *why* the underlying mathematics is true, and is part of an organized system of axioms, theorems, and results (Bell 1976; Coe & Ruthven, 1994; de Villiers, 1999; Harel & Sowder, 1998; Raman, 2003; Van Dormolen, 1977).

As an example of the third perspective on proof, Hanna & Jahnke (1993) defined proof as a "finite sequence of formulae within a given system, each formula being either an axiom or derivable from earlier formula by a rule in the system" (p. 423). Hersh (1993) agreed with this definition of a rigorous proof, but he stated that in reality, "proof is just a convincing argument, as judged by competent judges." Hersh continued by contrasting the two branches of mathematics: pure and applied. In pure mathematics (i.e., Hilbert's way of thinking), proofs are all about the rigor without regard to the usefulness of the result. In applied mathematics (i.e., Poincare's way of thinking), proof is all about the practicality of the result, which is based on compelling experimental evidence. Given the limited knowledge students have about the field of mathematics, it is unreasonable to expect that students will develop an abstract, axiomatic system in algebra, geometry, probability, statistics, and calculus during the time they are in high school. From this perspective, the purpose of proof should be to explain why results are true (Hersh, 1993) while giving students practice with deductive reasoning. In particular, students should develop some sort of *proof scheme*, which is the process for removing doubt about the truth of a statement for oneself (*ascertaining*) and removing the doubt for another person (*persuading*) (Harel & Sowder, 1998). How a student chooses to convince anyone about the truth of a claim depends on the student's conception of the field of mathematics (Harel & Sowder, 1998). "To construct a proof requires an essential shift in the learner's epistemological position: passing from a practical person (ruled by a kind of logic of practice) to a theoretical position (ruled by the intrinsic specificity of a theory)" (Balacheff, 2010, p. 118).

A. J. Stylianides (2007) took the main ideas from mathematicians' definitions of proof and framed them in a way that is useful for K-12 mathematics education. He defined proof as a "mathematical argument, a connected sequence of assertions for or against a mathematical claim, with the following characteristics (p. 291):

- It uses statements accepted by the classroom community (set of accepted statements) that are true and available without further justification;
- It employs forms of reasoning (modes of argumentation) that are valid and known to, or within the conceptual reach of, the classroom community; and
- it is communicated with forms of expression (modes of argument representation) that are appropriate and known to, or within the conceptual reach of, the classroom community"

His definition removes the apparent discontinuity in the transition from elementary to high school by referring to the "classroom community" and provides criteria that is accessible and instructive to teachers providing opportunities to their students to reason-and-prove. For the purposes of the study described herein, A.J. Stylianides' definition of proof will be used.

2.1.2 An Analytical Framework of Reasoning-and-Proving

G. J. Stylianides (2008, 2010) described an analytical framework (Figure 2.1) that captured many of the activities described by researchers involved in reasoning-and-proving. In these papers, Stylianides stated that he wanted to integrate reasoning with proving to remove the isolationist stigma attached to proofs in school mathematics. He also wanted to provide a tool for studying reasoning-and-proving activities among students. Stylianides argued that since new knowledge in mathematics frequently passes through four stages—identifying patterns, making conjectures, providing non-proof arguments, and providing proofs—students should be afforded the same scaffolding as they learn mathematics in school. The framework contains each of these stages. Stylianides termed the "overarching activity that encompasses these four activities" in an integrated way as *reasoning-and-proving* (p.9). Stylianides also argued that the three components (mathematical, psychological, and pedagogical) of the framework "can…provide the means to connect research findings from different investigations, thereby supporting the development of integrated knowledge across different domains" (Stylianides, 2008, p. 9).

<table>
<tr><td></td><td colspan="4">Reasoning-and-proving</td></tr>
<tr><td rowspan="3">Mathematical Component</td><td colspan="4">What are the major activities involved in reasoning-and-proving?</td></tr>
<tr><td colspan="2">Making generalizations</td><td colspan="2">Developing arguments</td></tr>
<tr><td>Identifying a pattern (plausible or definite)</td><td>Making a conjecture</td><td>Developing a proof (generic argument or demonstration)</td><td>Developing a non-proof argument (empirical argument or rationale)</td></tr>
<tr><td>Learner component</td><td colspan="4">What are the students' perceptions of the mathematical nature of a pattern /conjecture / proof / non-proof argument?</td></tr>
<tr><td>Pedagogical component</td><td colspan="4">How does the mathematical nature of a pattern / conjecture / proof / non-proof argument compare with students' perceptions of this nature?
How can teachers help their students reconsider and change (if necessary) their perceptions to better approximate the mathematical nature of a pattern / conjecture / proof / non-proof argument?

Stylianides, G. J. (2010). Engaging secondary students in reasoning-and-proving. Mathematics Teaching, September, 39-44.</td></tr>
</table>

Figure 2.1 G. J. Stylianides' analytic framework for the components and activities of reasoning-and-proving.

The mathematical component describes a hierarchy of reasoning-and-proving activities: identifying a pattern, making a conjecture, providing a non-proof argument, and providing a proof. The key to identifying a pattern is that students learn to see structural generalizations (linking patterns with features of the task) over empirical generalizations (just plausible patterns, no conviction for others). A conjecture is a reasoned hypothesis that has yet to be proven but extends the mathematics beyond the "domain of the cases that gave rise to it." Arguments that do not validate conjectures fall into two categories: empirical arguments and rationales. An empirical argument is based on the examination of a few tested cases that fit the proposed generalization with no attempt to extend the argument to cover all cases in a domain. A rationale is a partial proof; that is, an argument that is insufficient in some way, such as containing undefined terms or improper logical inferences. Using A. J. Stylianides (2007) definition of proof, G. J. Stylianides separates proof into two categories: generic examples and demonstrations. A generic example uses a particular case as an instance of the general case, while a demonstration uses formally established modes of mathematical proof, such as counterexamples, contradiction, and mathematical induction.

The psychological component of reasoning-and-proving pertains to the subjective nature of proof from a learner's perspective. For instance, intuition plays a part in the construction of a proof (Fischbein, 1999). It was Fischbein's opinion that intuition both helped and hindered students. While intuitions are necessary for discovering new strategies and models, the formalities in mathematical proof are so foreign to students that the process contradicts their natural intuitions. Fischbein described five possible situations in proof-making that are influenced by intuition:

1. a statement is so obvious that it is accepted without proof (e.g., two right angles make a straight line)
2. a statement is accepted but is also mathematically proven to support the intuition (e.g., vertical angles are congruent)

3. a statement is not intuitive and requires mathematical proof before acceptance (e.g., the sum of the angles in a triangle is 180 degrees)
4. a statement is intuitively accepted as probable but then is contradicted by a proof
5. two conflicting intuitions may appear

From a teaching perspective, "the requirement to prove a statement which appears [intuitively] obvious may strengthen the student's feeling that mathematics is an arbitrary, useless, whimsical game" (Fischbein, 1999, p. 22). Teachers, however, still need to encourage students to prove even seemingly obvious properties because properties do not always work the same way for every operation or context (e.g., commutative property holds for addition but not for subtraction). Fischbein recommended teachers helping students see that when intuition conflicts with formal mathematical evidence or proof, the formal proof is the authority, and since every aspect of mathematics does not lend itself to intuition (e.g., imaginary numbers), students need to accommodate the notion of mathematics as an abstract, deductive system of knowledge.

Balacheff (2010) tackled another aspect of the learner component in Stylianides' framework: the bridge between knowing and proving.

> The difficulty students may have [with trusting representations] relates not to their lack of mathematical knowledge but to a general human inclination not to question their knowledge and their environment unless there is a tangible contradiction between what is expected after a given action and what is obtained. (Balacheff, 2010, p. 123)

Balacheff described student knowledge and behavior in terms of adapting to an increasingly complex environment (formal mathematics). Initially, a student exists in equilibrium with his or her knowledge and understanding; the student does not encounter any conflicts with the mathematics being studied. Once a disturbance in the form of a contradiction or uncertainty is detected—throwing the student into cognitive disequilibrium—the student interacts with the environment within prescribed rules (perhaps by creating a proof) in order to return to a new state of safe equilibrium. Balacheff suggested that conceptions (instance of situated knowing) are validation-dependent and that the action of proving is a visible sign of the intellectual activity of validation.

The pedagogical component of reasoning-and-proving in Stylianides' analytic framework describes teachers teasing out student conceptions of proof and forming plans to align student conceptions with the field of mathematics. For example, teachers can help students "transition from knowing in action to knowing in discourse" (Balacheff, 2010, p. 131) by modeling conviction beyond solutions, giving status to students' arguments, and by helping students shift from the practical to the theoretical. According to Balacheff (2010), a claim has to be explicitly expressed and shared to even be considered for validation.

Another challenging aspect of reasoning-and-proving pedagogy is how to help students learn strategic knowledge. Weber (2001) investigated strategic knowledge by exploring the questions:

> Suppose we had a student who knew what a proof was, could reason logically, and was aware of and could apply the important facts, concepts, and theorems of a mathematical domain. Would this student necessarily be able to use that knowledge to construct proofs in that domain? If not, why not? (Weber, 2001, p. 102)

In other words, had this student developed intuition about reasoning-and-proving that suggested a path toward a proof? Weber found that when the undergraduate mathematics majors in his study failed to produce a proof, it was because they did not apply syntactic knowledge or recall basic facts. He found that students struggled when there was a large number of possible inferences or possible approaches to a proof. In contrast, the graduate mathematics students in Weber's study used their knowledge of a particular mathematical domain to increase their efficiency by choosing techniques commonly used in that domain, had a firmer grasp of important theorems, and were able to see when to use procedural or symbol manipulations and when to use conceptual or semantic knowledge.

In summary, as the field of mathematics evolved, so did the concept and role of proof, which influenced how mathematicians defined proof. A. J. Stylianides (2007) formulated a definition of proof that offers accessible and reasonable criteria for teachers that acknowledges their students (community of learners). G. J. Stylianides (2008, 2010) took the four scaffolded activities already used in the field of mathematics and three components important to mathematics education—mathematical, learner, and pedagogical—to devise an analytic framework for reasoning-and-proving that supports the development of knowledge for researchers, students, and teachers. In the next section, three aspects of teaching proof that challenge secondary teachers are discussed: knowledge base, insufficient resources, and unsupportive pedagogy.

2.2 ASPECTS OF TEACHING PROOF THAT CHALLENGE SECONDARY TEACHERS

There have been many empirical studies about secondary students, undergraduate students, and secondary teachers' conceptions of proof and ability to do proof (see Appendices A and B for summaries of empirical studies). The studies on practicing secondary teachers reveal that many teachers have a spotty knowledge of proof. On the successful side, most of the teachers in

Knuth's (2002c) study thought the role of proof in secondary school was to develop logical thinking skills, defined proof as a logical or deductive argument that demonstrates the truth of a claim, and believed that proof was a social act for communication. In evaluating proofs, the teachers choose mathematically grounded criteria and were able to correctly choose valid proofs with a success rate of 93% (Knuth, 2002a). Tabach, Levenson, Barkai, Tsamir, Tirosh, & Dreyfus (2011) found that practicing teachers had sufficient knowledge to construct proofs in elementary number theory. On the unsuccessful side, the studies revealed areas of weakness in the teachers' knowledge base and pedagogy.

2.2.1 Knowledge Base

In his landmark study on practicing secondary teachers' conceptions of proof, Knuth (2002b, 2002c) found that teachers had weaknesses with respect to understanding of the role of proof, concept of proof, rigor in proof, validation of proof, and ability to generalize. Knuth interviewed sixteen practicing secondary school mathematics teachers for his study. The teachers—who volunteered for the study—had a wide range of experience in the classroom (3 to 20 years, courses from pre-algebra to Advanced Placement calculus) and used a selection of reform-based and traditional curricula. Since the teachers sought reform-based professional development, Knuth assumed that the teachers were familiar with reform documents such as NCTM's *Principles and Standards for School Mathematics* (2000).

Knuth (2002c) compiled several author's views (e.g., Bell, 1976; de Villiers, 1999; Hanna, 1983; Schoenfeld, 1994) on the roles of proof to create a framework from which he designed his research questions. The framework described five roles of proof in mathematics (and school mathematics) (p. 63):

- To verify that a statement is true
- To explain why a statement is true
- To communicate mathematical knowledge
- To discover or create new mathematics, or
- To systematize statements into an axiomatic system

The practicing teachers in Knuth's study were interviewed twice. The first interview probed the teachers' conception of proof from the perspective of individuals who understand mathematics. The second interview probed the teachers' conceptions of proof as it pertained to secondary school mathematics. In the second interview, Knuth asked questions such as "what constitutes proof in secondary school mathematics? Why teach proof in secondary mathematics?" and "what do you think about the recommendations for proof set forth in the NCTM *Principles and Standards for School Mathematics?*" (p. 67). In order to clarify the teachers' responses, Knuth offered a series of constructed arguments—valid and invalid, explanatory and obtuse—and asked the teachers to describe whether or not they would use that particular argument "to convince students of a statement's truth" (p. 68). Tasks around which the arguments were constructed included: "Prove: The sum of the first n positive integers is $n(n+1)/2$" (2002b, p. 384) and prove that the sum of the angles in any triangle is 180 degrees (2002c).

Knuth (2002c) found that the teachers distinguished between "formal" proofs and "informal" proofs. The formal proofs were expected to use very specific language ("congruent" as opposed to "equal"), a specific format (e.g., two-column), and a particular method (e.g., proof by induction). The teachers accepted "less formal" proofs as long as the proof presented a sound, general, convincing argument, a view which was shared by Hersh (1993) (when the proof is also explanatory). The informal proofs consisted of non-rigorous explanations and empirical arguments. While the practicing teachers recognized these as invalid arguments in Knuth's (2002b) study, they did accept them as important stepping stones along the path from reasoning to proof and used specific examples as valuable illustrations of underlying mathematical concepts. Knuth (2002b) found that some teachers even judged some empirical arguments as acceptable proofs for high school students; 33% of the nonproof arguments shown to teachers were evaluated as "valid" proofs. What is not clear is whether the teachers regularly communicated to their students that non-proof arguments are instructive but insufficient. In light of the large body of research revealing students' preference and acceptance of empirical arguments (Chazan, 1993; Coe & Ruthven, 1994, Edwards, 1999; Goetting, 1995; Hadas, Hershkowitz, & Schwarz, 2000; Healy & Hoyles, 2000; Knuth, Choppin, Slaughter, & Sutherland, 2002; McCrone & Martin, 2009; Porteous, 1990; Schoenfeld, 1989; Stylianides, G. J. & Stylianides, A. J., 2009; Williams, 1979), students do not understand what constitutes a valid proof. One might imagine that if students hold external conviction proof schemes (authoritarian proof scheme, ritual proof scheme, or non-referential symbolic proof scheme) (Harel & Sowder, 2007) and their teacher tacitly or openly accepts empirical arguments, the student's conception of valid proofs will include empirical arguments (Chazan & Lueke, 2009).

The criteria practicing teachers used in evaluating proof-arguments included correct algebraic manipulations, sufficient detail, using an established method (e.g., proof by contradiction), and steps that were easy to follow (Knuth, 2002b). The arguments that teachers found the most convincing used a familiar method, showed insight into the underlying mathematics, proved the general case, and relied on specific examples or provided a visual reference. Through the interviews, Knuth realized that the teachers revealed some hesitancy in evaluating the logic of an argument, which could indicate a perceived weakness. The teachers instead focused on the correctness of symbolic manipulations or the selection of a valid proof method. If teachers tend to teach the way they were taught, examining the results of reasoning-and-proving studies on students might suggest why the practicing teachers have reasoning-and-proving knowledge base difficulties, such as with logic (see Table 2.1). For example, some undergraduate student studies point to students' difficulty with supporting warrants (Alcock & Weber, 2005; Inglis & Alcock, 2012; Ko & Knuth, 2009). Alcock and Weber (2005) reported the results of a study that looked at mathematics majors' ability to assess the logic in arguments, specifically, could the students distinguish between true statements and true statements that were warranted from previous statements in the proof? Alcock and Weber found that students were much better at catching unwarranted final statements in a proof than unwarranted statements within the body of the proof. In a different study of mathematics majors, Selden and Selden (2003) found that students with a "static" view of mathematics thought that mathematical competence implied knowing a large body of algorithms; in order to construct a proof, a student merely needs to select an appropriate algorithm and correctly implement the steps instead of building a deductive argument with logical validity. These studies indicate that practicing teachers and undergraduate mathematics majors have difficulty with evaluating the logic used in arguments, with the logical flow of arguments, and with the role of logic in proof.

Table 2.1 Summary of Empirical Studies: Difficulties in Knowledge Base across Subjects

	Students: secondary		Students: Undergraduate				Teachers	
Area of Difficulty	**6-8**	**9-12**	**non-specific**	**math majors**	**preservice elementary**	**preservice secondary**	**secondary**	**Empirical Studies**
Concept of Proof		X	X	X			X	Fischbein (1982), Knuth (2002c), Moore (1994), Selden & Selden (2003), Williams (1979)
Empirical Arguments	X	X	X		X	X	X	Chazan (1993), Coe & Ruthven (1994), Edwards, L. D. (1999), Gotting (1995), Hadas, Hershkowitz & Schwarz (2000), Harel & Rabin (2010), Knuth & Sutherland (2004), Martin & Harel (1989), Morris (2002), Schoenfeld (1989), Simon & Blume (1996), Stylianides, G. J. & Stylianides, A. J. (2009), Williams (1979)
Generalization	X	X		X	X	X	X	Chazan (1993), Ellis (2011), Fischbein (1982), Galbraith (1981), Knuth (2002b), Knuth & Sutherland (2004), Martin, et al, (2005), Morris (2002), Porteous (1990), Schoenfeld (1986), Selden & Selden (2003), Stylianides,

								Stylianides & Philippou (2007), Williams (1979)
Logic (chaining inferences)		X					X	Galbraith (1981), Knuth (2002b), Selden & Selden (2003), Senk (1985)
Logic (lack understanding)		X	X	X			X	Alcock & Weber (2005), Hadas, Hershkowitz & Schwarz (2000), Knuth (2002b), Ko & Knuth (2009), Selden & Selden (2003), Williams (1979)
Rigor	X	X					X	Fischbein (1982), Healy & Hoyles (2000), Knuth (2002b, 2002c), Kuchemann & Hoyles (2002), McCrone & Martin (2009)
Role of Proof (to understand)		X			X		X	Knuth (2002b), Mingus & Grassl (1999), Simon & Blume (1996), Tinto (1991)
Role of proof (to create an axiomatic system)		X						McCrone & Martin (2009), Tinto (1990)
Validation				X			X	Knuth (2002b), Selden & Selden (2003)
Validation (based on format)		X	X		X			Inglis & Alcock (2012), McCrone & Martin (2009), Stylianides, Stylianides & Philippou (2004)
Validation (authority?)		X			X			Edwards, L. D. (1999), Galbraith (1981), Simon & Blume (1996)

Another weakness Knuth (2002b) found in teachers was misunderstanding the generality of proof (the true power of deductive proofs). Morris (2002) and Selden and Selden (2003) conducted studies on undergraduate students that revealed problems with creating or understanding a proof that holds for all cases in a particular domain. Morris (2002) reported that students felt that since any axiomatic system is based on definitions (which may change) and postulates (which may or may not be true), an argument could not possibly prove the validity of a concept within a system. These students exhaustively checked cases even after they constructed a proof. The students seemed to think that there is ONE axiomatic system for ALL of mathematics, rather than separate axiomatic systems for different branches of mathematics. For example, in hyperbolic and spherical geometries, Euclid's Fifth Postulate (regarding a parallel line through a point not on another line) is false, but in Euclidean geometry, the Fifth Postulate is true. Students do not always realize that each geometry has its own axiomatic system, that statements which have been proven to hold for all cases in a domain are true and accepted within each system. Morris (2002) posited that the single axiomatic system view may be the case why 40% of the undergraduate students in her study accepted at least one inductive argument (note: 47% did reject all inductive arguments). Similarly, while ten practicing teachers in Knuth's (2002b) study demonstrated understanding of possible invalidity of a proof when moving to another axiomatic system, six teachers (almost a third) thought that even within a single axiomatic system, it might be possible to find a counterexample that would change a proof from valid to invalid.

2.2.1.1 Appropriateness of proof in secondary mathematics classrooms

Although many mathematicians and mathematics educators have called for proof to be a central idea throughout every level of school mathematics (Harel & Sowder, 1998; NGA Center & CCSSM, 2010; NCTM, 1989, 2000; Schoenfeld, 1994), empirical evidence suggests that some practicing secondary teachers disagree (Haggarty & Pepin, 2002; Knuth, 2002c). The teachers in Knuth's study thought that formal proof is beyond the grasp of many students and should only be taught to the most advanced students. One teacher stated that if she was asked to pare down the quantity of topics covered in her curriculum, she would jettison proof. Another teacher stated that he thought the authors of NCTM's *Principles and Standards for School Mathematics* (2000) were "smoking crack" (Knuth, 2002c, p. 75) when suggesting that all students learn to develop and evaluate proofs. Other teachers interpreted policy documents to fit their own beliefs: informal proofs were important and appropriate for all students but formal proofs were only appropriate for upper level mathematics students, and then perhaps restricted to Euclidean geometry courses (Knuth, 2002c). It is possible that these views were influenced by the challenges of teaching proof to all students. There is a strong connection between a teacher's beliefs and a teacher's practice (Schoenfeld, 1998). If a teacher does not believe that most students are capable of learning formal proof, the teacher will probably not spend much time teaching students how to construct valid arguments. Research indicates that secondary students do not get a strong foundation in proof (Hoyles, 1997). Several researchers described multiple conceptual errors in proof-making that could be the result of insufficient preparation and experience in secondary school. Selden & Selden (2003) found that undergraduate students made basic errors including beginning proofs with the conclusion, employing circular reasoning, leaving gaps in logic, making locally unintelligible arguments, and weakening the statement to be proven. Edwards and Ward (2004) and Moore (1994) described undergraduate students' difficulties with using definitions. Coe and Ruthven (1994) found undergraduate students unable to verbalize their thoughts, handle abstraction, and employ the context of the problem. All of these issues can be worked on in high school if a teacher believes them to be important and within the grasp of students.

In conclusion, teachers' own conceptual understanding of proof, their views on the role of proof and rigor, their understanding of generality, and their weak knowledge of logic describe a shaky knowledge base that effects teachers' ability to successfully engage themselves and their students in proof. One way to overcome some of these obstacles is through support offered to teachers via their curriculum. The next section describes several studies that investigated the reasoning-and-proving tasks offered in textbooks and the educative materials available to teachers in one textbook series.

2.2.2 Insufficient Resources: Textbooks

Teachers' reliance on their textbooks has been well documented in the United States, Sweden, Norway, Spain, England, and Germany (Battista & Clements, 2000; Grouws & Smith, 2000; Grouws, Smith, & Sztajn, 2004; Haggarty & Pepin, 2002; Horizon Research, Inc., 2003; Robitaille & Travers, 1992; Schmidt, McKnight, & Raizen, 1997; Tyson-Bernstein & Woodward, 1991). Tarr, Chavez, Reys, & Reys (2006) recently studied 39 teachers from eleven middle schools in the United States. Seventeen of the teachers used a reform-based curriculum and twenty-two of the teachers used traditional texts from major publishers. Tarr et al. found that regardless of curriculum type, teachers rarely added additional topics not covered by the textbooks and frequently culled some topics. 61-67% of the content the teachers chose was greatly influenced by the textbook and 41-52% of the ways the content was presented by the teachers was suggested by the textbook. Frequently skipped content focused on geometry, measurement, data analysis, and probability. Love and Pimm (1996, p. 402) wrote:

> The book is still by far the most pervasive technology to be found in use in mathematics classrooms. Because it is ubiquitous, the textbook has profoundly shaped our notion of mathematics and how it might be taught. By its use the of 'explanation-example-exercises' format, by the way in which it addresses both teacher and learner, in its linear sequence, in its very conception of techniques, results and theorems, the textbook has dominated both the perceptions and the practices of school mathematics.

When interviewed about their reliance on their textbooks, the teachers responded with comments such as "[The textbook] is like my bible. It is the basis for most of my instruction…source of homework, learning tool" (Tarr et al., 2006, p. 9), and "I use [the textbook] as a crutch—my curriculum. I'm letting [it] dictate what I teach" (Tarr et al., 2006, p. 8). These comments suggest that the teachers are allowing the textbook authors to make decisions regarding sequencing of material, content selection, available activities, and instructional techniques (Tarr et al., 2006).

Teachers' reliance on textbooks would not be problematic except that research also suggests that textbooks are insufficient resources for teachers who are attempting to transform their classrooms into standards-based environments. As early as 1983, the National Commission on Excellence in Education (NCEE) called for a reform of textbooks in their report *A Nation at Risk* (NCEE, 1983). The NCEE felt that publishers had responded to perceived market pressure to water down the reading level and content of their texts, and the NCEE called for more rigor and teacher input into the development of texts. In the next few decades, publishers attempted to meet the demands of individual state standards. The publishers gathered teacher feedback, revised texts, and sent samples to schools; however, all of these actions were expensive and hard to maintain (Reys, Reys, & Chavez, 2004; Seely, 2003). As a result, large states with approved-textbook lists, notably Texas and California, had disproportionate influence over the structure and content of the published textbooks (Seely, 2003). The market pressures eventually forced mergers of several publishing companies into three powerhouses: Holt McDougal, Pearson, and McGraw-Hill.

By this time, two versions of the NCTM Standards had been published (NCTM, 1989, 2000). The powerhouse publishers produced cross-walk guides that appeared to demonstrate the inclusion of NCTM content standards in their materials, but the fidelity was debatable. In an attempt to reach more teachers and students with the reforms suggested in the NCTM Standards, the National Science Foundation funded several reform-curriculum projects (e.g., the Connected Mathematics Project (CMP) and the Interactive Mathematics Project), some of which were purchased and distributed by one of the powerhouse publishers.

Regardless of publishers' efforts to produce standards-based curriculum (in reality or in appearance), research indicates that teachers filter the textbook author's intentions through their own belief systems and take-up or override reform suggestions (Knuth, 2002c; Lithner, 2004; Putnam, 1992; Remillard, 2005; Seely, 2003; Stein, Grover & Henningsen, 1996). As a result, there is a danger that an intended curriculum (topics and pedagogy) might not be implemented with fidelity due to teachers' filters, especially topics that barely surface, such as reasoning-and-proof? Content, organization, and sequencing of topics all contribute to students' conceptions of proof (Battista & Clements, 2000; Chazan, 1993; Healy & Hoyles, 2000; Hoyles, 1997), but without systematic attention to reasoning-and-proving, it is unlikely that students will understand mathematical proof or become proficient at it (Yackel & Hanna, 2003). In his study of mathematics teaching in eighteen high schools, Porter (1993) found that on average, no instructional time was devoted to offering students opportunities to reason-and-prove. The *exposure hypothesis* suggests that topics that receive the most exposure lead to the highest achievement (Mayer, Tajika, & Stanley, 1991). For example, elementary students in the United States are exposed to a higher ratio of language-based qualitative reasoning skills to computation-based reasoning skills than students in Japan. While the Japanese students showed higher achievement on computation-based questions, the U.S. students outperformed the Japanese students on problem-solving problems (Mayer et al., 1991). Thus, if students are not exposed to reasoning-and-proof, students will have little opportunity to develop any skills in this area. If textbooks—which have great influence on what is taught—do not attend to reasoning-and-proof, it is unlikely that students will receive the necessary exposure.

Even in situations where a textbook offers a great deal of exposure to reasoning-and-proving tasks and a teacher faithfully implements many tasks designed to offer students opportunities for reasoning-and-proving, the students' ability to interpret a text can pose problems. In general, students can approach a text two different ways: with a text-centered model (readers are receivers of meaning) or with a reader-centered model (readers are makers of meaning) (Weinberg & Wiesner, 2011). Reader-centered students are more likely to be successful at reasoning-and-proving because their learning characteristics include questioning, making connections, taking new perspectives, and constructing meanings socially (Borasi & Siegel, 1990). Text-centered students may not work to interpret the delimiters in proofs and thus fail to understand the logical arguments presented in proofs (Konior, 1993; Weinberg & Wiesner, 2011).

Recently there have been a few studies that created analytical methods to study reasoning-and-proving opportunities in curricula. The next section will discuss the coding and results from G. J. Stylianides (2009) study of the middle school Connected Math Project (CMP) curriculum, Thompson, Senk, and Johnson's (2012) study of high school mathematics texts from various publishers, and Lithner's (2004) study of undergraduate calculus texts.

2.2.2.1 G. J. Stylainides' (2009) study of the Connected Math Project

As discussed earlier, G. J. Stylianides (2008, 2010) designed an analytic framework that captured some of the major activities involved in reasoning-and-proving: making mathematical generalizations (identifying patterns and making generalizations) and providing support to mathematical claims (non-proof and proof arguments). Stylianides distinguished between two types of patterns: plausible and definite. A student might identify several *plausible* relationships between data, structures, properties, or variables. When the student selects one of the plausible patterns by comparing the patterns to the context of the problem, the pattern becomes *definite*. For example, the pattern suggested by $f(1) = 2, f(2) = 4$ could be modeled with the either the function $f(x) = 2^x$ or the function $f(x) = 2x$. Both functions are plausible. To become a definite pattern, the data could be linked to a particular context, such as cutting a piece of paper in half, stacking the pieces, and cuting in half again. Repeat and determine how many pieces result from each cut. The definite pattern associated with this process is $f(x) = 2^x$ (G. J.

Stylianides, 2009). The student's next step is to generalize the definite pattern by making a reasoned hypothesis (subject to testing and verification) extending beyond the data that was presented to the student. As with the patterns, these *conjectures* can be plausible or definite.

From the hypothesis, two types of arguments can be made: non-proof arguments and proof arguments. The weakest argument is an *empirical argument*, which is based on inductive reasoning. In this case, a student moves from the specific to the general; the evidence offered is validating that the conjecture holds true for a discrete collection of cases. Most arguments that are based on several examples are non-proof arguments but there are two exceptions: proof by exhaustion and proof by contradiction. Proof by contradiction only requires one example that shows the falseness of a conjecture. Proof by exhaustion, on the other hand, requires listing *every* possible case with each case supporting the conjecture. For instance, if young students are asked how many combinations of coins are possible given a penny, 2 nickels, and a dime, the students could actually generate every case, thus proving their answer is correct by exhausting all of the possibilities. In contrast, if an older student is asked to prove that the median of an isosceles triangle is always the perpendicular bisector from the same vertex, it would not be possible for the student to test every isosceles triangle because there are an infinite number of such triangles. Therefore, moving from an empirical argument into a proof by exhaustion is not possible in this case.

Stylianides called the next level of argument a *rationale*. A rationale is an argument that is almost a proof but is insufficient in some way, such as an argument that does not define the terms in the proof or back up statements with reasons. While rationales are not valid proofs, they do constitute an important type of activity in which students should engage because they represent good method choices, accessible thinking, and generally accurate conceptions of proofs. Frequently, students who pose rationales simply need a nudge in the correct direction to transform their arguments into proofs.

The ultimate reasoning-and-proving activity is constructing a proof argument. As defined by A. J. Stylianides (2007), "A *proof* is defined as a valid argument based on accepted truths for or against a mathematical claim that makes explicit reference to key accepted truths that it uses." G. J. Stylianides distinguished types of proof by "generic examples" and "demonstrations." A generic example is a specific case that is representative of the general case. For example, a student may generally represent an odd number as a series of paired dots with one left over without specifically mentioning the actual odd number the diagram represents. Stylianides suggested that students who have not learned specific, sophisticated proof methods can create generic examples to ascertain and persuade. The more sophisticated proofs are called demonstrations. Demonstrations are valid arguments by methods such as contradiction, exhaustion, and induction. It is difficult, however, for students to possess, access, and coordinate the content and proof knowledge required to construct demonstrations (Ko & Knuth, 2009, Moore, 1994, Weber, 2001). Sometimes students fall back on relying on heuristic methods (Raman, 2003) which work well for problem-solving but perhaps not so well for proof. Students relying on heuristic methods think that proof-making is about choosing the appropriate algorithm and faithfully following the steps of that algorithm (Selden & Selden, 2003). For example, Stylianides, A. J. , Stylianides, G. J., and Philippou (2004) and Stylianides, G. J., Stylianides, A. J., and Philippou (2007) found that preservice teachers in Cyprus had great difficulty with the methods of mathematics induction and contraposition, partly because the teachers did not fully understand the methods, as evidenced by their mistakes.

Stylianides (2007, 2009) applied his analytic framework to the opportunities designed in the Connected Mathematics Project (Lappan, Fey, Fitzgerald, Friel, & Philips, 1998/2004) for students and teachers to engage in reasoning-and-proving. Stylianides chose the CMP curriculum because as a reform-based curriculum, it was designed to "embody the recommendations of the NCTM (1989, 2000) Standards" (p. 260), and CMP was the most popular reform-based middle school mathematics textbook series in the United States when Stylianides worked on his study.

Stylianides coded tasks from the algebra, number theory, and geometry units in the CMP curricula from sixth, seventh, and eighth grade. He assessed every task along the actions required of the student and the purpose for that action. To make his decisions, Stylianides considered the knowledge (i.e. theorems, definitions, mathematical conventions, and methods) students should have learned by that unit and the approaches suggested by the student and teachers' editions of the textbook. For example, Figure 2.2 shows a task from the CMP unit *Looking for Pythagoras (*p. 44). Stylianides considered the predicted student answer in coding part b of the task. Since a student could answer the question correctly without needing to explain why the diagonal line bisects the corners of the square, the task only requires an answer to be a rationale, not a proof. Thus, part b of this task was coded as a rationale.

Square ABCD has sides of length 1. Draw a diagonal, dividing the square into two triangles. Cut out the square and fold it along the diagonal.

a.) How do the two triangles compare?
b.) What are the measures of the angles of one of the triangles? Explain how you found each measure.

Expected Formulation for part b (Teachers' Edition, p. 44):

The angle measures in each triangle are $45°$, $45°$, and $90°$. The diagonal line divides the corner angles into two equal angles, so the small angles must each be half of $90°$, or $45°$.

Figure 2.2 A Task from the CMP unit Looking for Pythagoras

Stylianides found that 38% of the tasks in the CMP algebra, number theory, and geometry units offered students some opportunities to reason-and-prove. Of these reasoning-and-proving tasks, 24% expected students to identify a pattern, 2% expected empirical arguments, 62% expected rationales, and 12% expected demonstrations (1% in seventh grade but 32% in eighth grade). There was a variety of reasoning-and-proving opportunities across content and grade levels. Students were offered the most opportunities for conjectures and proof in sixth grade, and the eighth grade curriculum contained the highest percentage of opportunities for students to create rationales. The fact that only 2% of the tasks were designed for empirical arguments speaks well for the CMP curriculum. Students may still provide empirical justifications for proof tasks, but these types of non-proof arguments are neither suggested nor encouraged by the textbook. Looking at this data by topic, number theory problems had the most number of making generalizations and proofs (see Table 2.2). Algebra had the fewest number of proofs, followed closely by geometry.

Table 2.2 Reasoning-and-proving Tasks in the CMP Number Theory, Algebra, and Geometry Units

Number of Opportunities	Identifying Patterns/ Making Conjectures		Non-Proof Arguments (Empirical / Rationales)		Proofs	
High	Number Theory	(16%)	Algebra	(27%)	Number Theory	(14%)
Medium	Algebra	(12%)	Geometry	(24%)	Geometry	(7%)
Low	Geometry	(11%)	Number Theory	(24%)	Algebra	(3%)

Stylianides summarized the four major findings of his CMP study according to desirable and undesirable features of a curriculum. The low percentage of tasks suggesting empirical arguments was desirable, but the low percentage of tasks requesting conjectures was undesirable. The high percentage of tasks designed to promote rationales was desirable, but the near absence of generic examples was undesirable.

2.2.2.2 Thompson, Senk, and Johnson's (2012) study of High School Mathematics Textbooks

Thompson, Senk, and Johnson (2012) took a broader look at reasoning-and-proving in curricular materials than did G. J. Stylianides (2009). Thompson, Senk, and Johnson called making and investigative conjectures, developing and evaluating deductive arguments, finding counterexamples, and correcting mistakes in logical arguments *proof-related reasoning* because these activities are "foundational elements of mathematical reasoning" (2012, p. 258). The researchers looked at exponential, logarithmic, and polynomial chapters in twenty contemporary high school mathematics textbooks. The topics were chosen because they contain a higher-than-average number of reasoning-and-proving tasks. Thompson, Senk, and Johnson further focused on exponential, logarithmic, and polynomial properties that are within the reach of high school students to justify. Both the narratives and the exercises of the textbooks were considered. For the narrations, Thompson, Senk, and Johnson coded how properties were justified. For the exercises, Thompson, Senk, and Johnson coded actions and whether the action was for a general or specific case. A list of their coding scheme and examples of each can be found in Appendices C and D.

Thompson, Senk, and Johnson examined 9742 tasks and found the overall percentage of tasks with reasoning-and-proving opportunities was very small (3.4% for Algebra 1, 5.2% for Algebra 2, and 7.6% for Precalculus). Thompson, Senk, and Johnson also aggregated the data by publisher and course. Separated by publisher, the results show that 14.7% of the exercises in Core Plus and 8.0% of the exercises in Key Curriculum Press offer students opportunities to practice reasoning-and-proving. The other publishers offer opportunities in only 3.5 – 6.3% of their tasks. Separated by course, the results show that students were offered about twice as many opportunities to reason-and-prove in precalculus courses as they were in the Algebra 1 course, and the opportunities in precalculus involved more general cases than the opportunities Algebra 1 which were largely focused on specific cases. The results of Thompson, Senk and Johnson's study indicated that textbooks did not off many opportunities for students to develop specific skills for reasoning-and-proving (see Table 2.3).

Table 2.3 Opportunities to Practice the Skills Involved in reasoning-and-proving in 20 Textbooks

Topic	% Tasks with Reasoning	Make Conjecture	Investigate Conjecture	Develop Argument	Evaluate Argument	Other	General/Specific
Exponents	5.1%	0.8%	1.6%	2.2%	0.1%	0.5%	1.7 / 3.4%
Logarithms	7.0%	0.6%	3.4%	2.5%	0%	0.7%	3.4 / 3.3%
Polynomials	5.0%	0.8%	1.1%	2.5%	0.1%	0.5%	2.8 / 1.9%
Overall	**5.4%**	**0.8%**	**1.7%**	**2.4%**	**0.1%**	**0.6%**	**2.6 / 2.5%**

2.2.2.3 Lithner's (2004) study of calculus books

Lithner (2004) conducted a study of reasoning in undergraduate calculus textbook exercises, using randomly chosen topics in textbooks from the publishers Addison-Wesley, Prentice Hall, and Wiley. The Addison-Wesley text was chosen because it had been used by subjects in a previous study and no reason was given for choosing the latter two publishers. While three textbooks constitute a small sample, Weinberg and Wiesner (2011) found that undergraduate calculus books have remarkably similar features, so the sample Lithner chose arguably represented the experiences offered to undergraduates in most calculus books. Lithner coded exercises according to how closely related the exercises were to examples in the narration of the text. The exercises with superficial reasoning allowed students to merely match surface features, achieving an answer by mimicking the procedures in a sample problem in the narration (IS). The next level of exercise also involved matching some procedural features to a worked problem but required the students to modify some components by employing local plausible reasoning (LPR). The exercises that required students to consider, understand, and use the intrinsic mathematical properties of the exercise were coded as *global plausible reasoning* (GPR). Table 2.4 lists the results for Lithner's study.

Table 2.4 Lithner's Results for Calculus Textbooks

Publisher: topics	IS	LPR	GPR
Addison-Wesley: limits and continuity	57%	29%	14%
Prentice Hall: polynomials and algebraic functions	85%	8%	7%
Wiley: applications of derivatives and integrals	56%	16%	28%

If a teacher assigned every exercise in Prentice Hall's polynomials and algebraic functions section, the students would merely be copying procedures in 85% of the tasks and engaging in novel reasoning in only 7% of the tasks. If a teacher assigned all of the problems in Wiley's applications of derivatives and integrals section, students would be forced to consider mathematics and reasoning in more than a quarter of the tasks. Of course, it is entirely possible for a teacher to assign no exercises that engage students in reasoning—over half of the tasks in each of the textbook sections Lithner considered only asked students to replicate procedures. As Selden and Selden (2003) claim, a large number of these superficial reasoning types of problems can lead students to a static view of mathematics, in which mastery of algorithms represents knowledge of the field.

Clearly, current textbooks do not generally offer students many opportunities to engage in Reasoning-and-Proving, regardless of whether the textbooks are traditional or reform-based, middle school or high school, Algebra 1 or calculus. If the exposure hypothesis (Mayer et al., 1991) is true, then neither teacher nor students will learn more about reasoning-and-proof through the resources at their disposal.

2.2.2.4 Educative Materials

Even if the opportunities to reason-and-prove are sparse in textbooks, research suggests that it is possible for curriculum materials to support teachers' learning in addition to student learning. In their seminal paper on the role of curriculum materials in teacher learning, Ball and Cohen (1996) called for textbook authors to consider curriculum *enactment* as well as curriculum content. The enactment of a lesson is influenced by the students, community, and school policy; by the teacher's understanding of the material, instructional design, and task selection; and by how the teacher foregrounds the intellectual and social environment of the class. Ball and Cohen suggested five ways that curriculum materials can help educate teachers by addressing these influences:

1. Help teachers prepare for the "unexpected" in the classroom by offering student artifacts which show a wide range of contextual student work
2. Help teachers learn content by highlighting multiple representations and their connections
3. Help teachers *hear* their students' ideas by discussing specific content revealed by

students' responses

4. Help teachers relate different units throughout the year by addressing the development of content
5. Help teachers plan for presentations by analyzing particular designs and considering students' understanding

Remillard (2000) also suggested providing samples of discourse and collections of nonstandard tasks for teachers to try. The major difference between educative materials and teacher guides is that educative materials help teachers develop general knowledge and learn to apply this knowledge flexibly in new situations (Davis & Krajcik, 2005). Nine years after Ball and Cohen (1996) published their work, Davis and Krajcik (2005) massaged Ball and Cohen's suggestions into a list of heuristics for developing educative curriculum materials (although Davis and Krajcik cautioned that their work needed further empirical testing). Each heuristic included three features: what the materials should provide to teachers, how the materials could reveal rationales for recommendations, and how teachers could take up those ideas.

The challenge of embedding educative material into curricula is to design the educative materials to meet the needs of teachers at a variety of levels, some of whom want a lot of detail and prescription and some of whom just need or want big ideas (Brown & Edelson, 2003; Davis & Krajcik, 2005). Educative materials that merely highlight or only address some of Ball & Cohen (1996) or Davis and Krajcik's (2005) recommendations may not be sufficient for teachers with impoverished understanding of proof. Educative curriculum materials that are too extensive run the risk of being ignored because teachers may not have the time to study the materials (Schneider & Krajcik, 2002).

Growing out of the concept of educative curriculum materials is the notion of *pedagogical design capacity.* This is the "teachers' ability to perceive and mobilize existing resources in order to craft instructional contexts" (Brown & Edelson, 2003, p. 6). While curriculum materials may influence and constrain teachers' actions, teachers also evaluate and enact curriculum through the lenses of their own experiences, intentions, and abilities (Brown & Edelson, 2003); it is virtually impossible to implement a curriculum with fidelity according to all of the author's intentions. "Teaching by design is not so much a conscious choice but an inevitable reality" (Brown & Edelson, 2003, p. 1). Thus, educative curriculum materials should be developed to promote teachers' pedagogical design capacity.

Theoretically, educative curriculum materials could help teachers design learning opportunities for their students to engage in reasoning-and-proving by helping teachers increase their pedagogical content knowledge and subject matter knowledge of proof (Yackel & Hanna, 2003). As defined by Shulman (1986), pedagogical content knowledge is an understanding of likely prior knowledge, misconceptions, and learning trajectories of students. G. J. Stylianides (2007) studied the CMP materials for guidance offered to teachers. It should be noted that CMP did not explicitly claim to contain educative materials that developed teachers' pedagogical design capacity, however, one of the program goals is to "help students and teachers develop mathematical knowledge, understanding, and skill" (Lappan et al., 2002, p. 1) and develop students' reasoning skills, meaning "bringing to any problem situation the disposition and ability to observe, experiment, analyze, abstract, induce, deduce, extend, generalize, relate, and manipulate in order to find solutions or prove conjectures involving interesting and important patterns" (Lappan et al., 2002, p. 6).

Stylianides (2007) evaluated every task in the CMP algebra, number theory, and geometry units for expected non-proof or proof arguments according to the approaches suggested by the student and teacher's editions and the student's expected knowledge to that point. Stylianides then separated the proof tasks into "solution only" and "solution with guidance" categories. The forms of additional guidance (educative materials) Stylianides considered were based in research (Ball & Cohen, 1996; Davis & Krajcik, 2005; Remillard, 2000, 2005; Stein & Kim, 2006) (Stylianides, 2007, pp. 197-198):

- Explanations about why students' engagement in a proof task matters
- Cautious points on how to manage student approaches to a proof task
- Discussions that support teachers' content knowledge of proof

Stylianides coded about 5% of the tasks in the algebra, number theory, and geometry units as proof tasks. Of these, only 10% were educative in some way, and not always in a positive way. For example, Stylianides described several issues with the forms of additional guidance attached to the topic of equivalence of symbolic expressions, from inaccurate use of mathematical language to issues with equivalence between different representations. However, this does not mean that the entire body of educative material was suspect, although the guidance offered was so little it was probably insufficient to help an impoverished teacher reform personal views and knowledge about proof.

2.2.3 Unsupportive Pedagogy

The research discussed so far indicates that some teachers' base knowledge and beliefs about reasoning-and-proving make it difficult for teachers to offer their students opportunities to develop their reasoning-and-proving skills and that some curricular materials provide insufficient models of justification, practice problems, and useful guidance. This section explores research about how teachers enact reasoning-and-proving activities in their classrooms. Some of the difficulties students and teachers have with proof can be addressed with supportive pedagogy or training (see Table 2.5). If teachers were using supportive pedagogy to help students develop their reasoning-and-proving knowledge, one would expect the trajectories of difficulties to be arrested between secondary students, undergraduate students, and teachers. This does not appear to be the case.   A small collection of studies (e.g., Ellis, 2011; Lampert, 1990; Mariotti, 2000; Yackel & Cobb, 1996) provide insight into how teachers might successfully break some of the cycles of difficulty with reasoning-and-proving. The first collection of studies—Lampert (1990), Mariotti (2000) and Yackel & Cobb (1996)—involves issues of authority and axiomatic systems, Ellis' (2011) study involves generalization, and Bieda's (2010) study involves discourse.

Table 2.5 Summary of Empirical Studies: Difficulties with Reasoning-and-Proving that can be addressed with Supportive Pedagogy

	Students: secondary		**Students: Undergraduate**				**Teachers**	
Area of Difficulty	**6-8**	**9-12**	**non-specific**	**math majors**	**preservice elementary**	**preservice secondary**	**secondary**	**Empirical Studies**
Context		X	X					Coe & Ruthven (1994), Küchemann & Hoyles (2001)
Definitions		X	X	X				Bell (1976), Edwards & Ward (2004), Moore (1994), Williams (1979)
Diagrams		X						Senk (1985)
Generalization	X	X		X	X	X	X	Chazan (1993), Ellis (2011), Fischbein (1982), Galbraith (1981), Knuth (2002b), Knuth & Sutherland (2004), Martin et al. (2005), Morris (2002), Porteous (1990), Schoenfeld (1986), Selden & Selden (2003), Stylianides, Stylianides & Philippou (2007), Williams (1979)
Launching		X	X					Moore (1994), Senk (1985)
Notation and Symbols			X	X			X	Moore (1994), Selden & Selden (2003), Tabach, et al. (2011)

	Students: secondary		Students: Undergraduate				Teachers	
Area of Difficulty	**6-8**	**9-12**	**non-specific**	**math majors**	**preservice elementary**	**preservice secondary**	**secondary**	**Empirical Studies**
Role of proof (to create an axiomatic system)		X						McCrone & Martin (2009), Tinto (1990)
Techniques (types of proof)		X	X	X	X	X		Chazan (1993), Knuth & Sutherland (2004), Ko & Knuth (2009), Stylianides, Stylianides & Philippou (2004, 2007), Weber (2001), Williams (1979)
Validation (authority?)		X			X			Edwards, L. D. (1999), Galbraith (1981), Simon & Blume (1996)

2.2.3.1 Pedagogical issues of sharing authority and conveying axiomatic systems

Moderating mathematically productive discourse is challenging, and common didactical teaching patterns will not bring empirical arguments closer to deductive arguments (Harel & Sowder, 1998; Hoyles, 1997; Weber, 2001). Chazan and Lueke (2009) found that when teachers follow a common demonstrate/practice instructional loop, they tend to guide students to the teachers' preferred method, act as the sole judge in disputes, and lecture with minimal student interaction. Suddenly asking students to explain why a process works makes the students question the validity of the process; it does not help the students develop reasoning skills. Teachers using these *authoritative proof schemes* (Harel & Rabin, 2010) in their classrooms do not encourage students to debate and resolve disagreements among themselves or probe their students' reasoning

Yackel and Cobb (1996) studied elementary classrooms to determine the factors that accounted for the development of mathematical beliefs and values in students. They pinpointed several aspects of classroom culture that influenced the beliefs and values which they termed *sociomathematical norms*. One of the goals of teaching reasoning-and-proof in classrooms is to foster the ability to convince and critique, so students need to understand "what counts as mathematically different, mathematically sophisticated, mathematically efficient, and mathematically elegant in a classroom" (Yackel & Cobb, 1996, p. 461). Yackel and Cobb found that when teachers asked their students for mathematically different solutions to problems, they effectively shifted authority away from themselves and towards students, thus empowering the students. While searching for mathematically different solutions, students participated in cognitively demanding work that allowed them to develop their reasoning ability. As students learned to use mathematics instead of status to validate arguments, the teachers' belief of the capabilities of students evolved. Yackel and Cobb described a reflexive relationship between students giving a "variety of explanations when different solutions are emphasized and developmentally sophisticated solutions are legitimized" (p. 466) and "teachers' evolving notions of what is sophisticated and efficient for the children" (p. 467). The students' explanations became objects of reflection themselves.

Lampert's (1990) description of her work with elementary school students is a good example of the benefits of setting sociomathematical norms. In her class, Lampert gave her students a sophisticated number theory problem to replace what could have been a procedure-driven lesson on exponents. In providing her students with a cognitively challenging task and removing herself as the sole mathematical authority in the room, Lampert gave her students opportunities to explore and negotiate meaning for themselves. Her students practiced convincing and persuading with an interesting task worthy of their time. In this way, Lampert's students used proof the way mathematicians use proof—to develop and extend knowledge.

Teachers' choices of activities and modes of discourse can impact students' mathematical beliefs and values (Martin et al., 2005; Yackel & Cobb, 1996). This is not to say that teachers do not have responsibility to lead students in developing their reasoning-and-proving skills (Balacheff, 1991). For example, Mariotti (2000) found that teachers can use discussion to foster their students' transition from intuitive to deductive reasoning. Mariotti conducted a design experiment over two years in Italy with secondary geometry students using Cabri geometry (a dynamic geometry environment) to develop an understanding of mathematical proof by developing an axiomatic system. During the year-long experiment, students were given figures to construct. In order to construct a figure whose features are preserved under “dragging tests”, the students had to understand the relationships among hierarchies of properties. For example, if a student was asked to construct a right triangle, the student could simply draw a segment, draw a connected segment that looked perpendicular to the first, and then connect the initial and final endpoints of the segments. This triangle would only look like a right triangle before dragging. A student who wanted to construct a right triangle whose 90 degree angle was preserved under dragging would have to pay attention to the order of the properties used in the construction. The student might construct an initial segment, then use a perpendicular command to construct a line perpendicular to the first segment through one of its endpoints. The student would truncate the line into a segment and connect the two free segment endpoints with the third segment of the triangle. Because the first angle was constructed with a perpendicular line, it will always remain a right angle, even if the triangle is stretched or moved. When the student records the steps and the reasons for the construction, the student is effectively writing a proof of why the triangle is a right triangle.

The secondary students in Mariotti’s (2000) study were given many such figures with specific and increasingly sophisticated properties to construct during class. While attempting to construct a figure, students were required to record the procedure they used and also why that procedure was correct. The procedure allowed students to impose an ordering on the sequence of axioms, definitions, and theorems they used, and collective revisions made during class discourse allowed students to connect and refine their thinking. The students’ classroom teacher helped students link constructing and dragging figures with conditional statements; in this way, students learned how to transform their empirical arguments into theoretical arguments. The teacher required students to justify their procedures according to rules which had been negotiated and accepted in class; the rules included previous constructions because each construction was shared, revised, and collectively accepted by the class. The teachers avoided authoritarian proof schemes by allowing their students to debate and resolve disagreements during collective revision time and by probing students' thinking during group construction time. Teachers further helped students develop an axiomatic system (Euclidean geometry) by providing the students with initial information (primitive entities, basic definitions, postulates) and requiring them to keep track of the negotiated and accepted rules (e.g., definitions and theorems) they subsequently developed. By the end of the course, the students had developed, recorded, and used the axiomatic system of Euclidean geometry.

2.2.3.2 Pedagogical issues with generalization

Ellis (2011) examined a different problem associated with reasoning-and-proving: generalization. While students may recognize a pattern, they struggle to use generalized language to express that pattern in an algebraically useful way (Chazan, 1993; Fischbein, 1982; Galbraith, 1981; Knuth & Sutherland, 2004; Martin et al., 2005; Porteous, 1990; Schoenfeld, 1986; Williams, 1979). Students tend to focus on creating data tables and identifying covariational patterns rather than on generalizing to the nth case (English & Warren, 1995; Pegg & Redden, 1990; Stacey & MacGregor, 1997; Szombathely & Szarvas, 1998). In her study, Ellis conducted a 15-day teaching experiment on a mixed-ability group of six 8^{th} graders. The students were interviewed both before and after a daily intervention during which the students changed the dimensions of a rectangle to study the effects on area and perimeter and link those changes to the parameter in the quadratic equation $y = ax^2$. Ellis wanted to know what types of student/teacher actions and interactions worked together to promote and foster generalizing. Ellis found seven categories of actions and interactions that resulted in generalizations much of the time (Ellis, 2011, p. 316):

- Publically generalizing (linking or expanding on an idea)
- Encouraging generalizing (prompting search for pattern beyond case)
- Encouraging sharing of a generalization or idea (formal or informal request)
- Publicly sharing a generalization or idea (revoicing or publically validating or rejecting idea)
- Encouraging justification or clarification (ask student to describe origins of idea, restate, or explain)
- Building on an idea or generalization (refine or use idea to create a new idea)
- Focusing attention on mathematical relationships (between features)

By using these actions and interactions, the teacher set the sociomathematical norms for reasoning-and-proving, and the students in the study shared ownership in creating a culture that encouraged justification in their classroom. The rich, cognitively demanding nature of the task chosen by the intervention teacher contributed to the opportunities students had to make multiple conjectures and justifications. The teacher also revised activities over the course of the experiment to accommodate the student-led nature of the inquiry (albeit teacher influenced). The set of the activities, actions, and interactions led Ellis to describe the type of generalization displayed by the students and teacher-researcher as *collective generalizing,* similar to *collective proving* (Blanton & Stylianou, 2002) and *collective abstraction* (Cobb, 2005 as cited in Ellis, 2011). In other words, for students to develop their generalization understanding and skills, a teacher needs to establish and promote public discourse in the classroom. To prepare teachers for this type of work, Ellis suggested professional development that "would foster teachers' abilities to (a) encourage justifying and clarifying, (b) publicly share students' contributions, (c) explicitly encourage generalizing, and (d) refocus students' attention on mathematical relationships" (Ellis, 2011, p. 336).

2.2.3.3 Pedagogical issues with discourse

Bieda (2010) conducted research similar to Ellis (2011), seeking to learn the nature of Middle School students' and teachers' actions and discourse during justifying and proving. Unlike Ellis, Bieda recruited participants from middle schools that had been using the CMP curriculum because the units contain opportunities for students to engage in reasoning-and-proving. Bieda referenced the results of G. J. Stylianides' CMP studies (2007, 2009), which indicated that 38% of the tasks in the algebra, number theory, and geometry units in CMP were written to offer reasoning-and-proving opportunities for students. Bieda selected participants from a middle school which had been using the CMP curriculum for six years. The seven teachers in the study (3 sixth grade, 2 seventh grade, and 2 eighth grade) had also participated in district-mandated professional development for CMP, which included daylong workshops to discuss research on mathematics education. Bieda noted that the experienced teachers were invested in promoting research-based, high quality mathematics education but had not focused their previous professional development on reasoning-and-proving. Still, with the selection of these teachers and this curriculum, Bieda had controlled for teachers' knowledge base, insufficient curriculum, and pedagogy (at least in part). With these major challenges minimized, would teachers be able to successfully implement reasoning-and-proving activities in their classrooms?

Bieda observed 49 investigations across the grade levels. All of the observed investigations contained tasks that offered students opportunities to reason-and-prove. The observations were supplemented by teachers' responses to pre- and post-observation interviews. The pre-observation writing prompts asked teachers about goals, task modifications, and ideal student responses. The post-observation interview questions asked teachers about the implementation of the investigation, about the teacher's responses to student thinking about reasoning-and-proving, and about the information that was "taken-as-shared" in the class. The answers to the pre- and post-observation questions allowed Bieda to check her interpretations of the class discourse.

Bieda found that teachers were able to help students make mathematical generalizations, but the teachers struggled to help students support their mathematical claims. Bieda found that teachers provided inadequate feedback: the feedback was neither sufficient to sustain discussions nor sufficient to establish criteria for proof. At times any justification was accepted, even for empirical arguments. As seen in Ellis' (2011) study, pressing students for generalized arguments is quite difficult. The teachers in Bieda's study did not devote enough class time to "establish standards for proof-related activity" (Bieda, 2010, p. 378), which resulted in a decrease of cognitive demand of the reasoning-and-proving learning opportunities. The teachers felt constrained by the limitations of their schedules and by trying to avoid authoritarian proof schemes. Regardless of constraints, teachers still need to set sociomathematical norms, and this includes expectations for justifications. Chazan and Ball (1999) suggested that teachers can stimulate, manage, and use disagreement as a source of "intellectual ferment" (a time for ideas to "bubble and effervesce" (p. 7)), during which teachers monitor and manage disagreement so that the argument remains centered on the mathematics and not on a particular student. If a disagreement spins out of control, students are unlikely to devote mental resources to reflection, which disrupts the goal of getting students to think deeply about the mathematics at hand and reconsider their mathematical positions.

In conclusion, even when teachers have a solid base of knowledge, use a reform-based curriculum with ample opportunities to reason-and-prove, and make good pedagogical decisions regarding authority and discourse, instruction intending to support students' development and understanding of proof is still challenging and can be superficial. "This suggests that greater emphasis is needed for middle school teacher preparation, professional development, and curricular support to make justifying and proving a routine part of middle school students' opportunities to learn" (Bieda, 2010, p. 380).

2.3 PROFESSIONAL DEVELOPMENT

2.3.1 Stability in the Education system

Teaching is a remarkably stable cultural activity, and teachers tend to teach the way they were taught (Ball, 1988b; Ball & Cohen, 1999; Stigler & Hiebert, 1999). When a teacher does take up an initiative—such as expecting students to listen and respond to each other's ideas—the initiative is difficult to sustain because the "education system" will try to restabilize itself (Ball, 1988b; Ball & Cohen, 1999; Stigler & Hiebert, 1999). It is not enough for a teacher to know subject matter, pedagogy, and students; a teacher also needs to be able to self-analyze teaching in order to consider change (Ball & Cohen, 1999; Little, 1993; Smith, 2001). If the teacher's vision of good teaching is not aligned with the desired vision of good teaching, the teacher needs to experience situations that contrast with the teacher's mental picture of good teaching (Ball & Cohen, 1999; Little, 1993; Loucks-Horsley & Matsumoto, 1999). The teacher needs to grapple with what it means to "know and understand mathematics, the kinds of tasks in which [her] students should be engaged, and finally, [her] own role in the classroom" (Smith, 2001, p. 3-4).

In order to destabilize a practice such as an authoritative proof scheme, teachers would need to be convinced that a new practice would provide better learning opportunities for students. *Transformative* professional development inspires changes in stable beliefs held by teachers about effective mathematical teaching and learning (Thompson & Zeuli, 1999). However, professional development is generally not designed to be transformative. Many administrators and policy-makers perceive professional development as a way to "update" teachers, so the learning opportunities presented to teachers tend to be discontinuous. Teachers are rarely offered a long-range professional development curriculum (Ball & Cohen, 1999; Little, 1993) that helps them learn in and from practice. Such piecemeal and limited professional development does not help teachers deepen their expertise and understanding of subject matter, both of which make a huge difference in how students learn (Loucks-Horsley & Matsumoto, 1999).

2.3.2 Transforming professional development

If reform is to happen in United States mathematics classrooms, professional development must be reformed as well. The focus should be on teaching—not teachers—to help students achieve learning goals (Stigler & Hiebert, 1999). Teachers need learning opportunities that help them study the big ideas of the discipline (Loucks-Horsley & Matsumoto, 1999), develop flexible understandings of the mathematics that they will teach (Ball & Cohen, 1999; Thompson & Thompson, 1996), practice making complex and subtle decisions about teaching (Little, 1993), learn self-monitoring and analysis (Ball & Cohen, 1999; Loucks-Horsley & Matsumoto, 1999; Stigler & Hiebert, 1999), and do all of this external to real classrooms in real time (Ball & Cohen, 1999). According to Ball and Cohen (1999):

> To learn anything relevant to performance, professionals need experience with the tasks and ways of thinking that are fundamental to the practice. Those experiences must be immediate enough to be compelling and vivid. To learn more than mere imitation or survival, such experiences also must be sufficiently distanced to be open to careful scrutiny, unpacking, reconstruction, and the like (p. 12).

The key is to study and learn from practice without being in the middle of practice.

In addition, the mathematics education research community has expressed interest in professional learning tasks, which are "complex tasks that create opportunities for teachers to ponder pedagogical problems and their potential solutions through processes of reflection, knowledge sharing, and knowledge building" (Silver, Clark, Ghousseini, Charalambous, & Seely, 2007, p. 262). These tasks employ artifacts of practice—what Smith (2001) calls "samples of authentic practice"—in the service of professional development (ICMI, 2004, as cited by Silver et al., 2007). Many effective professional development models use professional learning tasks and research frameworks, as well as guiding principles based on the frameworks and related tools (Boston & Smith, 2009).

Smith (2001) explains the reasons why samples of authentic practice, such as tasks used in a classroom, pieces of student work, and cases of teaching episodes help teachers transform their practice. Real artifacts can help teachers discuss abstract and complex ideas (such as generalizing), show the constant dilemmas of teaching (such as how to sequence student proof arguments during a whole class discussion), and reveal students' thinking about mathematics (such as whether or not empirical arguments are sufficient). When the teachers' talk is focused on authentic classroom experiences, the talk is less about personal opinion and preference and more about what signifies good teaching, what qualifies as acceptable work, and which responses show understanding (Ball & Cohen, 1999). Teachers can develop mathematical learning theories from examining these artifacts if they are taught to abstract generalities from particular situations (Smith, 2001).

A task chosen as a sample of authentic practice must be a task worthy of discussion; it must be a task from which teachers can advance their knowledge of mathematics and help them think about classroom practice. Stein and Lane (1996) found that cognitively-demanding tasks have more potential for student learning than low-level tasks; thus, cognitively-demanding tasks are worthy tasks for study in professional development. The teachers learn from solving tasks, discussing strategies and approaches, determining necessary prior knowledge, relating methods and representations, predicting residue, and determining how the task fits into their curriculum (Ball & Cohen, 1999; Doyle, 1988; Hiebert et al., 1997; Smith, 2001). As will be described in the next section, one focus of the Enhancing Secondary Mathematics Teacher Preparation program was transforming teachers' classroom practice with respect to learning opportunities afforded by cognitively-demanding tasks.

Analyzing actual student responses to tasks allows teachers to move beyond prediction of what students are likely to do on a given task—which is based on the beliefs of the teacher—to what students can actually do. This can reveal new information about strategies and misconceptions (Ball & Cohen, 1999; Carpenter, Fennema, & Franke, 1996; Stylianou & Smith, 2000). According to Smith (2001):

> Examining students' work can help teachers realize that children's ways of interpreting, representing, and solving problems are different from the teacher's, but their methods may be equally valid. In addition, it can help teachers develop the ability to interpret or make sense of students' solution strategies and forms of representations. (pp. 14-15)

As will be discussed shortly, the Cognitively Guided Instruction project used videos of elementary students solving problems. Because the students were young, video tapes of the students were used to capture the students' thinking and provided opportunities for rich discussion among the teachers in the study.

Case studies can help teachers see how to translate professional development into practice (Stein, Smith, Henningsen, & Silver, 2009). Carpenter et al. (1996) argued that teachers' informal knowledge of students' strategies and thinking is "not well organized, and it generally has not played a prominent role as teachers make instructional decisions" (p. 5). Teachers need frameworks to help them attend to aspects of students' thinking, interpret meaning, and incorporate their new understanding into their practice. According to Barnett (1991), teachers need contextualized knowledge (as seen in narrative cases) in order to implement changes in their complex mathematical education practice. Cases can be presented to teachers in either written or video form. Smith (2001) champions written cases over video cases because it allows the author to focus the reader's attention on a specific aspect of practice; other elements can be dropped or minimized. There is, however, evidence in the literature of teacher learning from both types of cases. For example, teachers in Barnett's (1991) study read cases focused on multiplication of fractions and then grappled with questions such as should discrete or continuous models be used to introduce multiplication of fractions? Another question revolved around the source of students' confusion: was it language or was the topic developmentally inappropriate? In a different study, Friel and Carboni (2000) used video-based cases to help preservice teachers see the learning opportunities offered to students in a student-centered classroom. The cases used by Friel and Carboni allowed the preservice teachers to focus on the students and the students' responses, which facilitated the preservice teachers' reconstruction of their beliefs (i.e., teaching does not have to be didactical). A third study—the BI:FOCAL project from the University of Michighan—that used narrative cases will be discussed in detail later in this section.

Ball (as cited in Smith, 2001) does provide some cautions about professional development employing samples of authentic practice. First, the professional development cannot be a stand-alone session; a curriculum needs to be formed around the samples. Second, samples should be chosen that are relevant and compelling but do not actually represent a teacher's personal situation. Finally, teachers need to look beyond the rich detail contained in the sample of practice and generalize what they are learning. Otherwise, professional development around samples of authentic practice becomes merely an exercise in analyzing.

2.3.3 Special Case: Preservice Teacher Education

According to Grossman, Hammerness, and McDonald (2009b), teachers need to possess more than a quiver of skills; they need to be "decision-makers and reflective practitioners" (p. 274). In this light, helping preservice teachers learn the core practices involved with making decisions and reflecting on practice is a necessity. Core practices are high-leverage actions, such as engaging students in choral counting, providing clear instructional explanations, leading classroom discussions, and developing professional relationships with colleagues and students. In general, core practices for preservice teachers are (Grossman et al., 2009b, p. 277):

- Practices that occur with high frequency in teaching;
- Practices that novices can enact in classrooms across different curricula or instructional approaches;
- Practices that novices can actually begin to master;
- Practices that allow novices to learn more about students and about teaching;
- Practices that preserve the integrity and complexity of teaching; and
- Practices that are research-based and have the potential to improve student achievement.

Currently, typical preservice teacher education programs offer a disjointed experience of learning knowledge, skills, and professional identity through separate foundation and method courses and separate settings (university and K-12 schools) (Grossman et al., 2009b). Grossman and her colleagues argued that preservice teachers should experience a curriculum that places core practices squarely in the center, a curriculum that decomposes the complex nature of teaching into focused, research-based bites of practice on which preservice teachers can deliberately focus on their attention. With the help of teacher educators, the preservice teachers can connect their accumulated knowledge into effective instructional routines. Preservice teachers can learn abstract theories, concrete skills, and professional identity in context through a fluid and connected process of learning to practice with conceptual and practical tools. In order to accomplish this shift in teacher education programming, Grossman recommends employing pedagogies of reflection and enactment.

Pedagogies of reflection involve learning opportunities that help teachers develop a purposeful and deliberate way for thinking through problems of practice (Jay & Johnson, 2002). According to Zeichner and Liston (1996, p. 20), "reflective teaching entails a recognition, examination, and rumination over the implications of one's beliefs, experiences, attitudes, knowledge, and values as well as the opportunities and constraints provided by the social conditions in which the teacher works." Teachers can and should reflect on how their practice will lead to change (Zeichner & Liston, 1996). Reflective teachers have the ability to focus on a single aspect of their pedagogy, use reframing and reflective listening to view that aspect from different perspectives, and discuss that aspect with other teachers in order to act on a "thorough and reflective understanding of events, alternatives, and ethics" (McKenna, 1999, as cited by Jay & Johnson, 2002). Jay and Johnson (2002) describe a framework which delineates dimensions of reflection (descriptive, comparative, and critical) and can be used as an instructional tool for preservice teachers. For instance, *comparative reflection* "reframe[s] the matter for reflection in light of alternative views, others' perspectives, research, etc;" preservice teachers can ask: "If there is a goal, what are some other ways of accomplishing it?" (Jay & Johson, 2002, p. 78). For example, a preservice teacher might be asked to reflect on possible sequencing of student work for whole-class discussion that has a high potential for student learning.

Pedagogies of enactment are learning "opportunities for novice teachers to 'practice' the various instructional routines that are central to core practices of teaching" (Grossman et al., 2009b, p. 283). These opportunities go beyond watching models of practice; they put the preservice teacher in the position of a decision-making teacher who can try a routine and receive feedback. For example, one part of preparing to orchestrate a classroom discussion is to anticipate student responses to a task. A preservice teacher can approximate this practice by self-generating all of the student solutions of which the teacher can think, "teach" the lesson to preservice peers to generate further solutions, then practice the lesson in a K-12 school placement and receive feedback from observers. In this model of teacher preparation, the teacher educator shifts roles from a provider of knowledge to a skilled instructional coach who offers opportunities to practice aspects of the complex practice of teaching and regular, specific feedback which allows novices to develop conceptual and practical tools.

In pedagogies of enactment, representations of practice are the ways the complex practice of teaching is portrayed to preservice teachers, including the focus of the representation and what that focus makes visible to novices. Examples of representations of practice include what Smith (2001) calls samples of authentic practice, such as student work and narrative accounts of an instructional episode (Grossman, Compton, Igra, Ronfeldt, Shahan, & Williamson, 2009a). A teacher educator can choose to show a video of a student working out a problem and ask preservice teachers to imagine that they are the student's teachers and to write assessing and advancing questions. In this way, the preservice teachers would not have to attend to other students, classroom management, or curriculum. The preservice teachers can focus solely on evidence of the student's thinking and on enacting the core practice of asking questions. The preservice teachers could watch the tape multiple times for various core practices because the episode is not happening in real time. When the teacher educator orchestrates a whole-class discussion about the preservice teachers' noticings about the student's mathematical understanding and their assessing and advancing questions, then the teacher educator would be offering the preservice teachers an opportunity to reflect. Because the teaching episode was not analyzed in real time, it was only an approximation of practice, but an appropriate one for helping preservice teachers develop conceptual and practical skills with minimal risk. While divorcing the learning opportunity from authentic practice in classrooms, the teacher educator can prioritize the skill acquisition of their preservice teachers, so when the preservice teachers transition from approximations of practice to authentic practice in classrooms, the preservice teachers will be prepared for success with a quiver of developed core practices.

Practicing teachers do not need *approximations* of classrooms because they are already immersed in their own classrooms. However, teachers need samples of *other teacher's classrooms* in order to dissect and analyze practice. With compelling and relevant samples of authentic practice, teachers—like preservice teachers—can develop deeper understandings of the mathematics they teach, analyze and practice making careful decisions, and learn self-monitoring. Three projects that used samples of authentic practice and pedagogies of reflection and enactment in professional development for teachers are discussed in detail in the remainder of this section. Each of the projects had results that connected the professional development to changes in teachers' practice. The projects are: Enhancing Secondary Mathematics Teaching Preparation Project at the University of Pittsburgh, Cognitively Guided Instruction at the University of Wisconsin-Madison, and the BI:FOCAL project at the University of Michigan.

Each project used a different type of sample of authentic practice to help transform teachers' practice.

2.3.4 Connecting changes in teachers' practice to professional development

2.3.4.1 Enhancing Secondary Mathematics Teacher Preparation (ESP) Project

The Enhancing Secondary Mathematics Teacher Preparation (ESP) project was a NSF-funded project designed to provide mentor teachers with professional development. The reason for the professional development was to provide consistent experiences of quality mathematics instruction between their preservice teachers' university and field experiences. The professional development was designed to support teachers' improvements in their instructional practices, develop their capacity to mentor beginning teachers and preservice teachers, and develop a shared vision of effective mathematics teaching consistent with the preservice teachers' university coursework (Boston & Smith, 2009).

The focus of the work was on cognitively-demanding tasks. The teachers were provided with a research framework, samples of authentic practice, and tools to guide their work in their classrooms. The research framework was the *Mathematical Tasks Framework* (MTF) (Figure 2.3), a framework that models the progression of a task from selection to implementation (Stein et al., 1996). Stein and her colleagues also developed the *Task Analysis Guide* (TAG) to help teachers identify the level of cognitive demand associated with a task (see Appendix E). The levels in the TAG are: memorization (lowest), procedures without connections, procedures with connections, and doing mathematics (highest). The tasks with the highest cognitive demand involve complex thinking and reasoning strategies, such as justifying reasoning to someone else by using proof by contradiction. No path is suggested by the problem, and the task focuses the student on the mathematical structure of the problem, not irrelevant surface features.

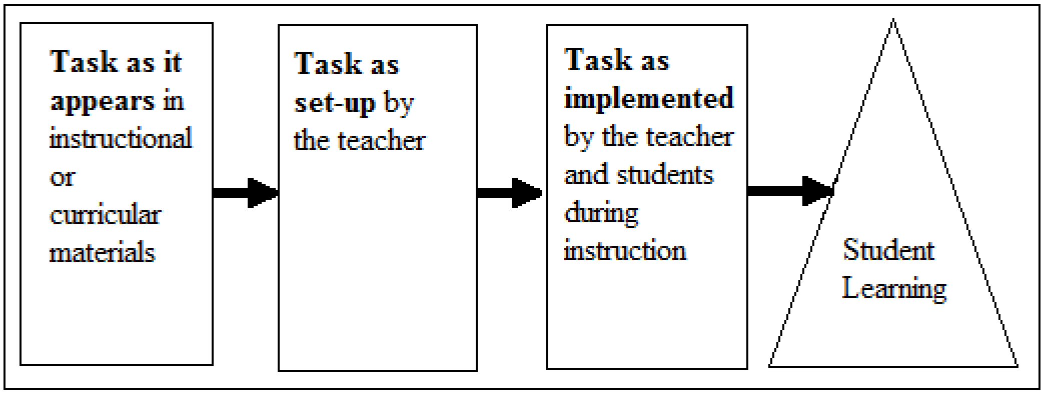


Figure 2.3 The Mathematical Tasks Framework (Stein et al., 2000).

The first year of the ESP project focused on improving the practice of the mentor teachers. The teachers analyzed samples of authentic practice, such as opportunities to solve, compare, discuss, adapt and categorize high-level tasks. They examined tasks with respect to the *Task Analysis Guide* (Stein et al., 1996), read and discussed cases, examined research about the factors and patterns of maintenance and decline of high-level tasks, analyzed student work, and wrote lesson plans. Throughout the training, the teachers reflected on their learning. In addition, teachers were provided with a lesson-planning tool.

To assess changes in teacher practice throughout the first year, Boston (2006) periodically collected classroom artifact packets (instructional tasks and student work) and performed related lesson observations. As hypothesized, the teachers showed an increase over time in their ability to select and implement cognitively-demanding instructional tasks. Boston argued that the high-level tasks offered by the teachers provided opportunities for students to develop their understanding of the underlying mathematics. This led to the teachers valuing high-level tasks and the impact high-level tasks have on students' learning more than before the professional development, which, in turn, led to the teachers selecting and implementing high-level tasks in their classrooms. This argument is in concert with Remillard's (2005) framework that suggests the participatory relationship between teachers and curriculum materials influences the planned curriculum and subsequently the enacted curriculum.

2.3.4.2 Cognitively Guided Instruction (CGI)

The CGI project was designed to help bridge research-based models of children's thinking and early elementary teacher's views of their own students' thinking. Carpenter and his colleagues hypothesized that providing teachers with knowledge about differences in problems, offering examples of students' strategies for solving those problems, and showing how those strategies evolve over time would directly affect teachers' classroom practice (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989). Carpenter, Fennema, and Franke (1996) later proposed that "understanding students' mathematical thinking can provide a unifying framework for the development of teachers' knowledge" (p. 4). The architects of CGI did this by asking teachers to reflect on and interpret the research-based models of students' thinking with respect to the teacher's own students and reflect on and interpret videos of students solving problems.

Carpenter et al., (1996) described a project based on early elementary students' intuitive knowledge about the four basic operations in mathematics: addition, subtraction, multiplication, and division, as seen in word problems. The CGI team created videos of students solving and explaining their processes for different types of required actions: joining action, separating action, part-part-whole relations, and comparison situations. The students' work illustrates different student strategies, such as counting all, counting on, and derived fact strategies. The CGI team also captured examples of students' solution strategies for multiplication, measurement-division, partitive division, rate problems, and array problems. The videos showed differences in strategies over time. "As they develop increasingly efficient ways to solve these problems, their understanding…increases concurrently with an understanding of how to apply this knowledge.…children can acquire the skills and concepts required to solve problems as they are solving the problems" (p. 10). In addition, some examples of students' work revealed combinations of strategies and relations between strategies. The collection of student work samples provided the basis for rich reflection among teachers about understanding students' intuitive strategies and developing understanding of whole-number concepts and operations, leading to a well-developed framework of understanding for the teachers. Teachers could then employ this "clearly delineated knowledge" (p. 16) in the classroom by offering students specific tasks to assess the students' understanding, to focus on different types of strategies, and to identify errors by listening carefully to the students' explanations of their solutions.

CGI studies (e.g., Carpenter et al., 1989) collectively indicate a positive correlation between considering students' thinking with changes in teachers' practice. The types of changes seen were incorporating more classroom time to problem-solving, expecting multiple solution strategies, and listening to children's explanations. The researchers also investigated the long-term impact of these changes (Fennema, Carpenter, Franke, Levi, Jacobs, & Empson, 1996). In this study, the teachers who participated in the CGI workshops were assessed with respect to their instruction and beliefs along a 5-level scale. The instruction scale ran from describing teachers who provided little opportunity to children to solve problems or share thinking (low) to teachers who provided many opportunities and whose instruction was driven by knowledge of individual students and was adapted based on what was shared. The belief scales ran from teachers believing that direct instruction was necessary (low) to teachers who believed that children can solve problems without direct instruction and that the teacher's understanding about children's thinking should drive her decision making. After the CGI training, almost every teacher changed the bulk of their instructional time to various problem-solving opportunities. The teachers were not bound to their textbooks, did not emphasize rote procedures, and spent more time listening to students than talking. A similar change was seen levels of beliefs. Virtually all of the teachers showed positive change in beliefs (children could and would learn mathematics by solving problems and discussing solutions). There was, however, no consistency in which change drove the other: did a change in beliefs cause a change in instruction or was it the other way around? It was also not clear why there was variety in the amount of change in teachers' instruction and beliefs. What was clear, though, was that an increase in students' performance was directly related to changes in teachers' instruction.

2.3.4.3 Silver and Mills: BI:FOCAL Project

The Beyond Implementation: Focusing on Challenge And Learning (BI:FOCAL) project was designed to offer middle school mathematics teachers practiced-based professional development (Silver, et al., 2007). The problem addressed in the BI:FOCAL project was that teachers implementing an innovative curriculum steeped in cognitively-demanding tasks sometimes hit a "curriculum implementation plateau," which is when trained teachers become comfortable with some new practices but the professional development stops before the teachers have adopted all of the new practices advocated in the innovative curriculum (Silver, Mills, Castro, & Ghousseini, 2006). The BI:FOCAL project was designed to help teachers fill in these gaps (mathematical knowledge, use of student thinking, and proficiency with pedagogical strategies) to become maximally effective.

The twelve teachers who participated in the BI:FOCAL project were teaching with the CMP curriculum (Lappan, Fey, Fitzgerald, Friel, & Philips, 1996) and met monthly for a day of professional learning tasks, consisting of modified lesson study and narrative case analysis and discussion. The lesson study helped teachers attend to general instructional goals and issues, asked teachers to adopt an analytical stance towards teaching in general, and treated an instructional episode as a unit of analysis (Silver et al., 2006). These activities connected aspects of teacher knowledge (e.g., understanding how students learn) to each other and to classroom practice. The narrative case analysis made classroom practice public, asked teachers to adopt an analytic stance towards their practice, and treated an instructional episode as a unit of improvement. The narrative cases (developed originally under the auspices of the Qualitative Understanding: Amplifying Student Achievement and Reasoning project and the Cases of Mathematics Instruction to Enhance Teaching project) used in the professional learning tasks were designed to focus the teachers' attention on particular, important aspects of mathematics teaching and learning. The narrative cases described the implementation of a task and showed the interactions among the students, task, and teacher during an instructional episode. Not only did the narrative cases help teachers see the connections between content and practice, but the cases helped teachers develop their understanding of mathematics education ideas: problem-solving, pedagogical dilemmas, student learning, and planning. The teachers were asked to solve a challenging task, read a narrative case illustrating the implementation of a similar task, discuss the case, and then reflect on their own practice as they collaboratively planned lessons. The teachers were thus afforded learning opportunities from three perspectives: as a mathematician solving problems, as an observer of teaching and learning, and as a facilitator of student learning. Each activity with its perspective had the potential to draw out misconceptions, impoverished understanding, or opposing viewpoints. As such, there were many opportunities for teachers' learning and consequently a transformation of their practice.

To assess the impact of the BI:FOCAL project, Silver and his colleagues (2007) collected and analyzed videos of the professional development sessions, written reflections of the participants, and semi-structured interviews with participants. Some of the ideas teachers noticed from a narrative case were the case subject's encouraging multiple representations, using questioning strategies, highlighting connections among solutions, and the sequencing of student solutions during whole-group discussions. Similar to the ESP and GCI projects, the teachers participating in BI:FOCAL's professional learning tasks were presented with such a full experience of exploration of these pedagogical ideas that the teachers could organize their thinking about content, pedagogy, and student learning into a framework that would help them be more effective teachers.

The results of the analysis of the project indicated that teachers had opportunities to build connections among mathematical ideas, connect these ideas to their practice, and consider a range of pedagogical actions and decisions with respect to student opportunities to learn. The narrative case analysis "generated a number of insights that provoked [teachers] to modify their instruction" (Silver, Ghousseini, Gosen, Charalambous, & Strawhun, 2005, p. 299). Silver et al. found evidence of transformation in the teachers' practice: using multiple solutions as a pedagogical tool to help advance the mathematical instructional goals of a lesson. The narrative cases had allowed teachers to unpack this nuanced, complex practice and view it as a means to make connections between mathematical ideas.

The three projects described in this section all sought to help teachers make connections between teachers' actions and student learning: use of high-level tasks in the ESP project, use of frameworks to guide problem selection and discussion in the CGI project, and using an analytical stance towards teaching in the BI:FOCAL project. Each project employed samples of authentic practice to engage teachers in transformative professional development, including narrative and video cases, cognitively-demanding tasks and samples of student work, research-based frameworks, and modified lesson study. The result of each project was a change in teachers' beliefs about the inherent potential of students and what learning opportunities can maximize that potential. In summary, transformative professional development is needed to help teachers overcome obstacles to providing opportunities for students to learn, including opportunities for students to learn reasoning-and-proving. Targeted professional development should be steeped in authentic teacher practice, utilizing research, case studies, samples of students' work, and an opportunity to grapple with content.

2.4 EXAMINING PRACTICE

2.4.1 Ways to investigate instructional practices

Stein, Baxter & Leinhardt (1990, p. 640) argued that in order to make the link between subject-matter knowledge and teachers' instructional practice, educational researchers "need to develop and draw upon detailed, qualitative descriptions of how teachers know, understand, and communicate their subject matters." This line of research coincided nicely with the publication of the first NCTM Standards (1989) which advocated developing connections, flexibly using multiple representations, and communicating discoveries which would lead to student competencies such as solving complex, multi-step problems. Since the NCTM Standards advocated teacher practices that were vastly different than how teachers were trained (Ball, 1988b), professional development was essential in helping teachers develop their subject-matter knowledge so that they could support their students' learning. The stream of research advocated by Stein et al. (1990) would help inform the type of professional development that was needed.

Case studies of individual teachers allow researchers to gather extensive amounts of data and perform a fine-grained analysis of the instructional practice of those teachers. For instance, the Stein et al. (1990) study described the subject-matter knowledge of a seasoned fifth grade teacher in an urban school by examining the teacher's practice via videotaping instructional episodes, interviewing, and gathering background knowledge. Stein and her colleagues defined subject-matter knowledge as "a combination of beliefs and knowledge about both mathematical content and content-specific pedagogy" (1990, p. 642). For this particular study, the mathematics content was functions and graphing. To assess the teacher's mathematical knowledge base, a card sort task was given to the teacher and a mathematics educator and the results were compared. The cards contained samples of representations, mathematical relationships and definitions of functions. The videotaped lessons served to reveal how the teacher's subject-matter knowledge influenced the teacher's presentation of main concepts, procedures, and relationships.

The researchers found that the teacher's limited subject-matter knowledge affected his instruction by not offering enough groundwork to his students for future lessons, overemphasizing concepts that were only partially correct (e.g., definition of a function), and missing opportunities to help students see connections among representations and concepts. Consequently, the teacher's actions implied a goal of factual fluency over conceptual fluency for functions, which was counter to the goals stated in reform documents (NCTM, 1989).

Case studies of individual teachers provide deep exploratory information but are are time-consuming, labor-intensive, and not necessarily generalizable. A different avenue for researchers to study instructional practices is analyzing the information gathered from a survey. While surveys are cost-effective and useful for gathering coarse, generalizable contrasts and similarities among teachers in large systems, surveys can contain inaccurate responses due to teachers misunderstanding terms and described situations described. For instance, a survey might ask teachers if they provide instruction on proof-making. If a teacher has an impoverished understanding of the core elements of proof, the teacher might respond that he or she is "teaching" proof without actually doing so. Classroom artifact packets, in contrast, provide more detailed information than surveys without the cost of observing teachers for case studies, making them an ideal vehicle for examining practice at scale.

2.4.2 Classroom artifact packets

Classroom artifact packets, as described by Boston (2006), include an instructional task, a description of the instructions given to students and the structure of the work, and samples of student work labeled high, medium, and low according to the teacher's evaluation criteria. The task, student work, and labels inform researchers about the teachers' values, expectations for student learning, and the cognitive processes for which students are held accountable. In the remainder of this section, three studies will be described that explored the use of classroom artifact packets: Clare and Aschbacher (2001), Borko, Stecher, Alonzo, Moncure, and McClam (2005), and Boston (2006).

2.4.2.1 Clare and Aschbacher (2001)

Clare and Aschbacher (2001) investigated using teachers' assignments (tasks) to gather information about teachers' practice in the classroom. Clare and Aschbacher argued that the collection and analysis of teacher's assignments were more cost-effective than observing practice directly, would yield more accurate results than surveys, and were better indicators of the learning opportunities teachers gave students to produce high-quality work. In their 2001 study, Clare and Aschbacher examined and reported the technical quality of using teacher assignments to assess instructional quality.

To create the rubrics needed to assess the teacher assignments, Clare and Aschbacher drew on research which indicated that teachers who offered high-quality learning opportunities to their students were knowledgeable about their content, set clear instructional goals, aligned instructional goals with lessons and assessments, focused on metacognitive strategies in their curriculum, communicated their expectations, provided feedback, and provided opportunities for students to share their knowledge. Six rubrics were then created: cognitive challenge of the task, clarity of the learning goals, clarity of the grading criteria, alignment of goals and task, alignment of goals and grading criteria, and overall quality. The created rubrics were used to assess the match of instructional quality between teachers' assignments with classroom observations for 32 third and seventh grade language-arts teachers working in an urban district. The teachers submitted four assignments, each containing an information cover sheet, four samples of what they deemed to be high-quality student work, and four samples of medium-quality work. Two classroom observations and interviews were made to allow comparison with the classroom artifact packets.

After training raters with anchor papers and other samples, Clare and Aschbacher had three raters score each assignment according to the rubrics (the reported inter-rater reliability was acceptable at about 83%). The student work was scored by two raters using established rubrics on organization, content, and mechanics. The classroom observations were made by trained and seasoned researchers and graduate students; the instructional quality of each lesson was assessed according to the cognitive challenge of the lesson activities, the quality of the classroom discussions, the level of student participation in classroom discussions, the quality of instructional feedback, the level of student engagement in the lesson, the lesson implementation/classroom management, the clarity of the learning goals, and the alignment of goals and lesson activities. Clare and Aschbacher found that the teacher's observation scores from the quality of classroom discussions rubric, the level of student participation in classroom discussion rubric, the quality of instructional feedback, and the cognitive challenge of the lesson activities correlated well with the quality of the teacher's assignments. The researchers also found a significant correlation between the quality of student work and the quality of teacher's assignments. In addition, most of the variation in the scores was due to actual differences across teachers, not raters. On the other hand, the teacher's observation scores from the level of student engagement rubric, the clarity of the learning goals, and the alignment of the goals and lesson activities were not significantly correlated with the quality of the teacher's assignments. In addition, reading comprehension and writing assignments were more correlated with the rubric scores than content-area writing assignments and challenging major projects. Clare and Aschbacher also found that the rubrics produced scores that were significantly correlated with each other, indicating that some rubrics could be eliminated, but not which ones. The researchers did not test for the influence of the quality of the teacher's assignments on the quality of the students' work, nor did they claim their results—from a small sample of urban districts—could be generalized.

2.4.2.2 Borko et al. (2005)

Borko and her colleagues (2005) addressed two of the unresolved issues of Clare and Aschbacher's (2001) study: generalizability and type of assignment correlation. Borko et al. looked mathematics and science to address generalizability and gathered all of the student work for an assignment (instead of just four samples) to address the type of assignment correlation. Thirteen middle school math and science teachers from California and Colorado participated in the study. The teachers taught in different types of schools (urban, suburban, and rural) and used either reform curriculum or traditional curriculum.

Each teacher submitted artifact packets representing their classroom practice for one week and each teacher was observed for 2-3 days during the same week. The five middle school science teachers constituted the pilot study. The classroom packets contained artifacts that represented a scoop of a teacher's practice: instructional materials, assignments, quizzes and tests, student work, feedback, wall displays, lesson plans, and teacher reflections. Teachers were directed to separate the artifacts into three categories: materials created in preparation for class (e.g., scoring rubrics), created during class (e.g., student work), and created after class (e.g., student homework). They were also directed to submit samples of high and medium quality student work.

The researchers used the *National Science Education Standards* (National Research Council, 1996), the *Principles and Standards for School Mathematics* (NCTM, 2000), and the experts in the Mosaic II project (Stecher et al., 2002 as cited in Borko et al., 2005) as guides in developing the scoring rubrics for quality science and mathematics teaching. The resulting science dimensions were: collaborative grouping, materials, assessment, scientific discourse, structure of instruction, hands-on, minds-on, cognitive depth, and inquiry, plus an overall rating. The resulting mathematics dimensions were: collaborative grouping, structure on instruction, multiple representations, hands-on, cognitive depth, mathematical communication, explanation and justification, problem solving, assessment, and connections-applications, plus an overall rating. The teacher quality score from the classroom observations was determined by applying the scoring rubrics to the observer's field notes. To increase inter-rater reliability, each observer was trained on videotaped lessons; discussion regarding the lesson and scores occurred until consensus was reached. The classroom artifact packets were scored by a team of researchers, both those who collected the data and discussed the scoring rubrics and the packets and those who had not collected data (to check the validity of scoring an artifact packet without the benefit of classroom observation). Unlike the method used in the Clare and Aschbacher (2001) study, average rater scores of teaching quality from the classroom practice artifacts were *combined* with the teacher quality scores of the classroom observation ratings and the interview questions to create the truest, most complete picture of teaching quality possible.

Borko and her colleagues found that the inter-rater reliability for the science classroom artifacts packets was “substantially higher than one would expect by chance alone” (2005, p. 86). As in the Clare and Aschbacher (2001) study, Borko et al. found that the rater agreement varied across the dimensions; the agreement was good for the dimensions minds-on, cognitive depth, inquiry, and overall, but not as good for materials (sufficient quantities of materials allowed access to information and support investigations) and assessment (formative). Interestingly, Borko and her colleagues noted that classroom artifact packets from traditional classrooms had a lower score (as expected for the rubric dimensions being used in the study) but a higher inter-rater reliability than packets from reform classrooms. Similar results were found for the mathematics artifact packets.

The correlation in scores of teaching quality between the mathematics artifact packets was moderate and varied across dimensions (higher for hands-on, cognitive depth, and overall ratings and lower for collaborative grouping and assessment) but not type of instruction; the differences in correlation was not related to reform or traditional curriculum. However, there was a difference in average score of instructional quality between reform (higher) and traditional (lower) curriculum. In addition, Borko and her colleagues found that mathematical practices that are difficult to capture with artifact packets—notably connections/applications and structure of instruction—had low inter-rater reliability when the classroom packets were considered by themselves. Subsequent discussions indicated that information about school context was necessary to determine whether the classroom instruction helped students connect what they were learning in mathematics with their environment. Similarly, when the artifact packet score in isolation was compared to the packet/observation/interview score, the dimensions collaborative grouping and multiple representations had lower reliability than the other dimensions.

In general, though, Borko and her colleagues found that teacher quality can be reasonably assessed via classroom artifact packets, regardless if the classroom content was science or math, the type of school was urban, suburban, or rural, and the type of instruction was reform or traditional. Inconsistencies in ratings were attributed to raters' inferences when presented with incomplete information or mixed messages from teachers transforming their practice from traditional to reform curricula, or to the ability of classroom artifact packets to describe interactive practices of teaching such as classroom discourse, or disagreement of what reform-oriented classroom practices look like.

2.4.2.3 Boston (2012)

While Borko and her colleagues looked mathematics and science to address generalizability and type of assignment correlation, Lindsay (Clare) Matsumura continued her work on rubrics to assess instructional quality through observations and student work samples. By 2006, Matsumura and her team had developed the Instructional Quality Assessment (IQA) (Matsumura, Slater, Junker, Peterson, Boson, Steele, & Resnick, 2006). The same year, team members Boston and Wolf (2006) published a paper that described the development and field testing of the IQA for mathematics, called the IQA Academic Rigor in Mathematics rubric (IQA AR:Math). The IQA Mathematics rubrics indicate the quality of teaching and student learning opportunities based on the teacher practices of selecting and implementing cognitively challenging tasks, providing opportunities for students to share their mathematical thinking, and sharing expectations for students' learning. The rubrics—which evaluate the cognitive demands in each practice—are based on the Mathematics Tasks Framework and the Task Analysis Guide from the QUASAR project (Stein, et al., 1996). Because all of the evaluations look at some form of cognitive demand, comparisons across the teaching practices can be made, providing researchers with a picture of the quality of instruction in a particular classroom. In the study conducted by Boston and Wolf (2006), inter-rater reliability was consistent (60-67.3%) for all four measures of student work samples (classroom artifact packets). The IQA AR:Math rubrics showed significant differences in mathematics instruction and learning between teachers who had received professional development around academic rigor, accountable talk, clear expectations, and self-management of learning and teachers who had not received training. The mathematics instruction and learning included understanding, sense-making, and the use of multiple representations. This initial study also reinforced Clare and Aschbacher's (2001) finding that classroom artifact packets can be used as stable indicators of classroom practice.

Over time, the collection of rubrics designed to assess instructional quality in language arts and mathematics was refined and renamed as the Instructional Quality Assessment (IQA) Toolkit (Boston & Wolf, 2006; Matsumura et al., 2006). Some of the rubrics are designed to be used for lesson observations and some for student work sample packets. The rubrics for mathematics specifically evaluate the quality of instruction based on four indicators and include: the Potential of the Task rubric for the cognitive demands of the task prior to instruction, the Implementation of the Task rubric for the students' engagement and thinking evident in the majority of students during the lesson, the Rigor in Students' Written Responses rubric for the use of multiple representations, etc., and a collection of three rubrics (Rigor of Teacher's Expectations, Clarity and Detail of Expectations, and Student Access) for teachers' expectations for students' learning and are based on information provided by the teacher, such as directions given to the student (Boston, 2012). The evaluation of the quality of instruction from the IQA Mathematics Toolkit was found to correlate well with student achievement (Matsumura, Garnier, Slater, & Boston, 2008; Quint, Akey, Rappaport, & Willner, 2007).

The recent study published by Boston describes her efforts to determine the "aspects of teachers' ability to enact high-quality mathematics instruction [that] are captured by the IQA Mathematics rubrics for lesson observations and collections of students' work" (p. 77). Boston collected data during the 2004-2005 school year on 13 middle school teachers who had participated in professional development in preparation for teaching a standards-based curriculum. The professional development encouraged the use of cognitively-demanding tasks and provided support for task implementation and classroom discourse. Boston collected 26 lesson observations (problem-solving lessons with class discussion around students' work) and 35 classroom artifact packets which contained challenging problem-solving tasks, six samples of student work, description on cover sheet. The IQA Mathematics Toolkit was used to assess the lesson observations and the classroom artifact packets. To ensure reliability, each lesson was observed by two raters and the classroom artifact packets were independently scored by three raters who resolved discrepancies with consensus.

The results of the lesson observations indicated that the trained teachers struggled with selecting cognitively-demanding tasks and implementing them at a high-level (median score of 2 out of 4 for Potential of the Task and Implementation of the Task). The classroom mathematical discussions were found to be weak as well; while many students participated, there was very little rigor, linking or press in the discussion around mathematical understanding. Only 27% of the lessons were considered to be high-level. Not surprisingly, the Rigor of Teacher's Expectations rubric indicated that teachers generally expected students to engage at the level of memorization or procedures without connections to meanings and no further.

The classroom artifact packets were not for the lessons that were observed and evaluated. For these separate assignments, the IQA Mathematics results were much different. The written assignments showed a higher quality of instruction; the rubrics for Potential of the Task, Implementation of the Task, Rigor of Students' Responses, and the collection for Teachers' Expectations had a median scores of 3 out of 4. 80% of the assignments were considered to be high-level. Thus, it appeared from the results of the IQA Mathematics Toolkit that teachers were better able to engage students in high-level mathematical thinking with written assignments than with classroom discussion. Boston conjectured that "supporting students to produce high-quality written work may be less pedagogically demanding than maintaining students' high-level engagement throughout an instructional episode" (pp. 96-97).

Note that Boston was not studying the correlation between classroom artifact packets and classroom observations for measuring instructional quality. Evidence of that correlation can be found in one of Matsumura's studies. Matsumura and her colleagues (2008) found that the IQA Mathematics Toolkit results from classroom observations and classroom artifact packets to be highly correlated for their study.

Since the IQA Mathematics rubrics are designed to only capture and assess certain instructional practices, Boston cautioned that the rubrics "are best suited for assessing reform-oriented instructional practices for use in implementation studies or curriculum or professional development, or to identify changes in the nature of school- or district-wide instructional practice over time (Boston, 2012, pp. 96-97). Other types of instructional practices (e.g., direct instruction) would not receive high scores on the IQA Mathematics rubrics. Boston also cautions that student work may not capture certain instructional moves, such reducing the cognitive demand of the task by smoothing the curriculum. The lesson observation rubrics capture these moves well, but not the student work rubrics.

In summary, a valid and reliable examination of a mathematics teacher's practice can be made through classroom artifact packets. These student work sample packets allow for the examination of classroom practice at scale because it is less time consuming than observations and more reliable than surveys.

This chapter described literature from four areas of mathematics research: the nature of proof, aspects of proof that challenge secondary teachers, professional development, and examining practice. Recent sets of mathematics content and practice standards (NGO Center & CCSSM, 2010; NCTM, 1989, 2000) captured the current views on the nature of proof, and the Stylianides have provided us with an accessible definition of proof for secondary students (A. J. Stylianides, 2007) and a scaffolded framework of reasoning-and-proving activities (G. J. Stylianides, 2008, 2010). Regardless of these standards, definitions, and framework, teaching reasoning-and-proof remains a complex activity due to teachers' weak knowledge base, insufficient resources, and unsupportive pedagogy. Transformative professional development which uses authentic artifacts of practice (e.g., tasks, student work, narrative cases) has been shown to help teachers improve their practice. In particular, teachers who have had this type of transformative professional development are capable of selecting, implementing and assessing the work from cognitively-demanding tasks.

The current study seeks to understand this transformative process with a focus on reasoning-and-proving. Do teachers who have participated in professional development with a focus on reasoning-and-proving (a subset of cognitively-demanding tasks) enact reasoning-and-proving tasks more often and more skillfully than the teachers described in the literature who did not have similar opportunities? In the next chapter, the research questions, data, coding, and analysis is described that seek to answer this broad question.

3.0 METHODOLOGY

The goal of this study was to determine how secondary teachers who have experienced targeted professional development enact reasoning-and-proving in their classrooms. This study investigated the ability of teachers who participated in a reasoning-and-proving course to select, implement, and evaluate the student products of reasoning-and-proving tasks. The teachers' abilities were examined through answers to background interview questions, logs of tasks used in class, textbook exercise analysis, and classroom artifact packets.

Because this study was about a particular phenomenon (enacting reasoning-and-proving) in a specific context (the teachers' classrooms), a qualitative research design was appropriate (Bogdan & Biklen, 2007; Erickson, 1986). The nature of the study was descriptive rather than causal and was bounded by four cases; therefore, the specific type of qualitative study was a descriptive multiple-case design (Merriam, 1998; Miles & Huberman, 1994; Yin, 1993). Since all of the teachers in the study received a similar professional development (CORP course), Yin (1993) suggested the study follow a replication logic, which is when consistent results over multiple cases are considered to be robust findings with confidence.

My intention was not to generalize how all teachers enact reasoning-and-proving in their classrooms, but to offer information about how these particular teachers selected, enhanced, implemented, and evaluated tasks in order to offer their students reasoning-and-proving learning opportunities. The choices the teachers made should have revealed what they understood and valued about reasoning-and-proving and how they integrated mathematical reasoning-and-proving into their curriculum. This study highlighted the complex nature of enacting reasoning-and-proving in secondary classrooms, especially for early-career teachers. As such, it was not my intention to evaluate the teachers in this study. It is my hope that this study contributes to our understanding of how trained teachers offered reasoning-and-proving learning opportunities by the way of tasks.

Stein et al.'s (1996) Mathematical Tasks Framework provides direction in selecting, launching, and implementing cognitively-demanding mathematical tasks. Tasks that provide students opportunities to develop their understanding of reasoning-and-proving are cognitively-demanding tasks (Stein & Smith, 1998; G. J. Stylianides, 2010). Since students come to understand the field of mathematics by working on cognitively-demanding tasks (Doyle 1983, 1988), studying the extent to which teachers select, implement, and evaluate the products of tasks after focused professional development was an appropriate lens through which to investigate the teachers' ability to engage students in reasoning-and-proving. The research questions were:

1.) **To what extent did participants select reasoning-and-proving learning opportunities in the form of tasks?**
    a.) To what extent does the textbook include tasks that have the potential to engage students in reasoning-and-proving?
    b.) To what extent did the participants select tasks for instruction that have the potential to engage students in reasoning-and-proving?

c.) To what extent did the participant modify tasks to affect the tasks' potential to engage students in reasoning-and-proving?

d.) What were the sources of the tasks that participants selected for instruction?

2.) To what extent were participants able to maintain the level of cognitive demand of the reasoning-and-proving task during implementation?

3.) To what extent were participants able to accurately evaluate their students' reasoning-and-proving products?

a.) To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?

b.) To what extent did participants apply the core elements of proof in evaluating their students' reasoning-and-proving products?

c.) In what ways did participants communicate expectations regarding what was required to produce a proof to students?

The following sections detail the reasoning-and-proving course in which the teachers participated, what data was be collected, how the data was be coded, and how the data was analyzed. The first section describes the context of the reasoning-and-proving course in which the teachers participated and a description of those teachers who participated in this research study. The next section explains the data collected in the study, specifically tasks (available, selected, enhanced, and sources), classroom artifact packets, and pre-and post- interview questions. The third section details the coding systems and methods of analyzing the data as related to the research questions. Validity and generalizability are briefly discussed in the final section.

3.1 PROFESSIONAL DEVELOPMENT: THE REASONING-AND-PROVING COURSE

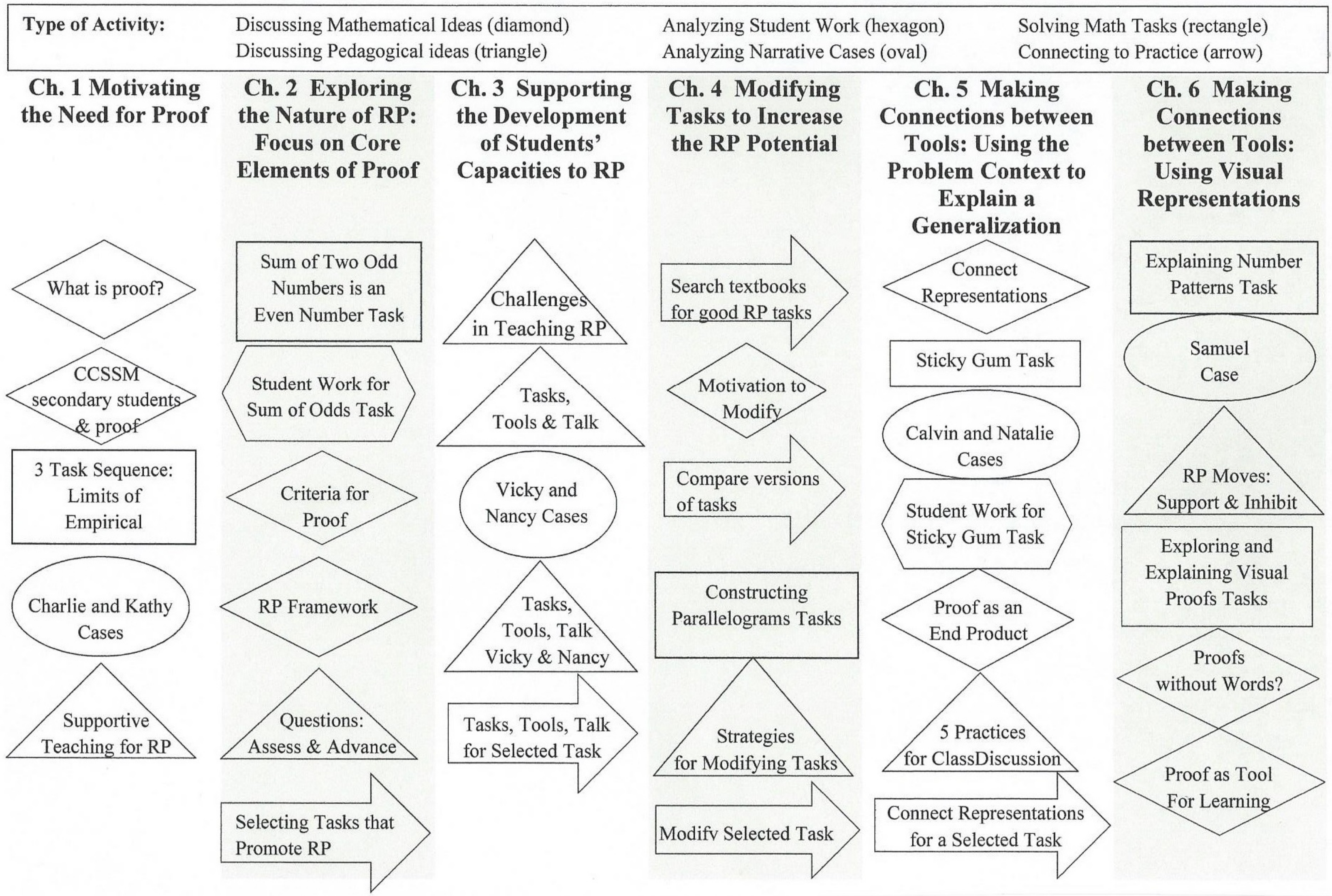
The goal of the NSF-funded Cases of Reasoning-and-Proving (CORP) project was to create professional development materials that would develop teachers' *knowledge of mathematics for teaching* (KMT) reasoning-and-proving, such as the ability to modify tasks or make connections between representations. The specific KMT contained in the CORP materials came out of the research that revealed obstacles to implementation, insufficient of resources, and the pedagogical difficulties of high school teachers offering reasoning-and-proving learning opportunities to their students. Common obstacles to implementation are: weak concept of proof, ability to generalize, understand rigor in proof, understand the role of proof, misunderstanding format (Knuth, 2002a, 2002b) and preference and acceptance of empirical arguments (e.g., Chazan, 1993). Although teachers rely heavily on their textbooks (Love & Pimm, 1996), studies have shown that the number of reasoning-and-proving tasks available in textbooks is small (Lithner, 2004; G. J. Stylianides, 2009; Thompson, Senk, & Johnson, 2012). Studies have also revealed three issues with teachers' pedagogy: misplaced authority (Lampert, 1990; Yackel & Cobb, 1996), understanding generalization (Ellis, 2011), and discourse that lowers the cognitive demands of reasoning-and-proving tasks (Bieda, 2010). The CORP materials address all of these issues. The topics contained in the CORP materials are: motivating the need for proof (limitations of empirical arguments), exploring the nature of proof with a particular focus on the core elements of proof, supporting the development of students' capacities to reason-and-prove through tasks, tools, and talk, modifying tasks to increase their reasoning-and-proving potential, and making connections between tools (using problem contexts and using visual representations).

The intended outcomes of the CORP course for teachers were an understanding of what constitutes reasoning-and-proving, an understanding of how secondary students benefit from engaging in reasoning-and-proving, and an understanding of how they can support the development of students' capacities to reason-and-prove. Initial evidence from pilots conducted at three large research universities suggest that teachers engaged with the materials, believe that what they learned in the course had the potential to impact their teaching practice, and that the cases helped teachers think about instructional issues related to reasoning-and-proving. In addition, the tasks-tools-talk structure was a useful framework for supporting teachers' analysis and discussion of the cases (Smith, Arbaugh, Steele, Boyle, Fulderson, Knouck, & Vrabel, 2012).

3.1.1 Course Activities

The delivery of the CORP content was grounded in elements of teachers' practice, as suggested by Ball & Cohen (1999). Table 3.1 lists the specific activities in which the participants engaged related to the major topics of the course. The participants in the professional development discussed the mathematical ideas, explored the pedagogical ideas, solved tasks, analyzed student work on tasks, analyzed narrative cases, and connected all of this work to their teaching practice. In preparation for discussing mathematical ideas, the participants read published research articles on the topics: analytical framework for reasoning-and-proving (G. J. Stylianides, 2010), tasks, tools, and talk (Chapin, Anderson & O'Conner, 2003; Hiebert et al., 1997), reasoning-and-proving tasks in high school textbooks (Johnson, Thompson & Senk, 2010), the value of connecting representations (Clements, 2004), and proof as a tool to learn mathematics (Knuth, 2002a). The student work and cases contained common errors and opportunities to practice writing proofs; the discussions, presentations, and reflections helped participants develop their pedagogy and correct misunderstandings regarding reasoning-and-proving.

Table 3.1 CORP Course Participant Activities by Chapter



In the beginning of the course (column one in Table 3.1), participants explored the mathematical concept of proof, how to recognize when a claim has been proven, and how proof can be studied in the context of education. The participants then addressed one of the common misconceptions about proof-making: empirical arguments do not constitute valid proofs but deductive arguments do constitute valid proofs (e.g., Chazan, 1993). The participants worked through a three-task sequence chosen by G. J. Stylianides and A. J. Stylianides (2009) to force participants who had accepted empirical arguments into cognitive conflict, which helped them see the limits of inductive reasoning.

A main goal of the second chapter (column 2 in Table 3.1) was to have participants develop criteria for proof which contained the core elements of a valid proof: the argument must show that the conjecture or claim is (or is not) true for *all* cases, the statements and definitions that are used in the argument must be true and accepted by the community because they have been previously justified, and the conclusion that is reached from the set of statements must follow logically from the argument made (based on A. J. Stylianides, 2007). Additionally, the participants needed to become aware that the type of proof, form of the proof, the representation used, and the explanatory power of the proof can vary for a valid proof. The participants arrived at these criteria by examining samples of student work around the task "prove that the sum of two odd numbers is even." The samples of student work varied with respect to format, level of sophistication, and validity; the variety provided opportunities for the participants to determine what is necessary and what was optional for a valid proof. The CORP materials provided the course instructor with sample questions and suggestions about massaging the participants' criteria into a focused list. After examining student work on different tasks throughout the course, participants were asked to refine their criteria. By the end of the course, participants had practiced using a refined criteria to evaluate proof products that contained the core elements of proof.

Another goal of the second chapter was to have participants explore different reasoning-and-proving activities. While current curriculum standards advocate infusing mathematics curriculum with reasoning (NGA Center & CCSSO, 2010; NCTM, 2000), specific types of reasoning activities help students develop their mathematical proving skills (Lakatos, 1976; G. J. Stylianides, 2008). These activities-- searching for patterns, making conjectures, and providing non-proof and proof arguments—are reflected in G. J. Stylianides' (2010) analytical framework for reasoning-and-proving (see Figure 2.1). Stylianides argued that mathematicians process new knowledge through the four stages of the framework and students should be afforded the same scaffolding in school. His framework contains each of these stages. Stylianides also argued that the three components of the framework (mathematical, psychological, and pedagogical) "can…provide the means to connect research findings from different investigations, thereby supporting the development of integrated knowledge across different domains" (Stylianides, 2008, p. 9).

The participants next focused on three major aspects of teaching, namely tasks, tools, and talk. In introducing the task, tools, and talk lens in Chapter 3 (third column in Table 3.1), the architects of the CORP materials provided a concrete way for the participants to consider, anticipate, and incorporate reasoning-and-proving into their classrooms. The participants were first asked to anticipate the challenges of integrating reasoning-and-proving into their practice, such as communicating valid criteria for proof to their students and fostering a classroom culture that supported students' development of reasoning-and-proving skills. Next the participants read, discussed, and summarized readings on tasks, tools, and talk (Chapin, Anderson &

O’Conner, 2003; Hiebert et al., 1997). The two narrative cases—*The Case of Vicky Mansfield* and *The Case of Nancy Edwards*—asked the participants to tease out different ways teachers can support their students’ understanding of reasoning-and-proof. The participants then combined their observations into a list of pedagogical moves that might address the anticipated classroom challenges, a list to which they added throughout the rest of the course.

Throughout the course, participants wrote reflections. In Chapter 3, the participants wrote about selecting tasks that allowed all students to develop their understanding of reasoning-and-proving and about implementing those tasks with a focus on tools and talk. All the participants were preservice teachers who had already studied the Math Tasks Framework (Stein et al., 1996), levels of cognitive demand of mathematical tasks as described in the Task Analysis Guide (Stein & Smith, 1998), and the Five Practices for Orchestrating Mathematical Discussions (Stein, Engle, Smith, & Hughes, 2008) in previous courses. Participants brought this information as well as the course readings on tasks, tools, and talk into their reflection responses. For example, some participants noticed that the teachers in the Chapter 3 cases selected good tasks (Task Analysis Guide), fostered collaboration by allowing students to think and talk, continually pressed students for justification, sequenced student presentations of solutions (three of the Five Practices), and launched the task effectively (Math Tasks Framework).

The fourth chapter of the course (column four in Table 3.1) addressed the lack of good reasoning-and-proving tasks in published curriculum material (Lithner, 2004; G. Stylianides, 2009; Thompson, Senk, and Johnson, 2012). Since many teachers rely heavily on their textbooks (Battista & Clements, 2000; Grouws & Smith, 2000; Grouws, Smith, & Sztajn, 2004; Haggarty & Pepin, 2002; Horizon Research, Inc., 2003; Robitaille & Travers, 1992; Schmidt, McKnight, & Raizen, 1997; Tyson-Bernstein & Woodward, 1991) the teachers need to learn to modify tasks in order to increase learning opportunities for reasoning-and-proving for their students. After the teachers searched their own textbooks for good reasoning-and-proving tasks, the teachers read about Johnson, Thompson, and Senk's (2010) efforts to identify reasoning-and-proving tasks and property justifications in many popular, commercially-available textbooks. Armed with very specific knowledge of verbs that encourage reasoning-and-proving actions from the article they read (e.g., *write a convincing* argument, *determine* the error in reasoning), the participants then explored ways to modify insufficient tasks. After examining several sets of tasks and suggested modifications, the participants created a list of strategies for modifying tasks to increase their reasoning-and-proving potential. Teachers then practiced using these strategies to modify additional tasks.

The final two chapters of the CORP materials (columns five and six in Table 3.1) focused on using the problem context to generalize a solution and the explanatory power of visual proofs. The material in these chapters, while important, was not directly related to the current study.

3.1.2 The Participants

In the summers of 2011 and 2012, 10 and 8 preservice teachers, respectively, participated in the CORP course on reasoning-and-proving as part of their teacher certification program. From these 18 preservice teachers, four agreed to participate in this study. All of the participants in the CORP course were asked to participate in the study, but not every teacher had secured a full-time teaching job at the time of the study, which effectively shrunk the pool of potential participants. A few others who did secure teaching positions were concerned about taking on any additional responsibilities so early in their careers. Five teachers originally agreed to participate; one was removed due to the timing of the study. Two of the study participants experienced the professional development in 2011 and completed a full year of teaching prior to the study. The other two study participants experienced the professional development in 2012 and were in their first year of teaching during the study (see Table 3.2). The value in asking early-career teachers to participate in a study in which an important mathematical topic (reasoning-and-proving) is poorly represented in textbooks but is largely represented in curriculum standards is that early career teachers tend to pay more attention to planning and have fewer instructional routines (Sleep, 2009) than experienced teachers, and can thus better accommodate new ideas into their curriculum.

Table 3.2 The Participants in the Current Study

Participant Name	School	# of Yrs. Teaching	Subject	Curriculum Materials
Sidney	Suburban Middle School in Virginia	1	Algebra 1	Prentice Hall *Algebra 1*, VA edition
Jonathan	Urban Charter School in Pennsylvania	1	Algebra 1	Prentice Hall *Algebra 1*, Common Core edition
Karen	Urban Magnet High School in Pennsylvania	2	Geometry	CME *Geometry*
Uma	Urban High School in Virginia	2	Geometry	Glencoe *Geometry*, VA edition

3.2 DATA COLLECTION

Four data sources were used to answer the research questions: the class textbook or curriculum guide, task log sheets of relevant tasks, classroom artifact packets, and pre- and post-data collection interviews. The class textbook or curriculum guide indicated which tasks were available to the teacher and students. The task log sheets indicated what tasks were selected, which tasks were modified and how, the sources of the tasks, and the purpose of the tasks. The classroom artifact packets provided information on how students engaged with the task and about how the teacher evaluated students' reasoning-and-proving products. The background (pre-) interview questions intended to capture the freedom the participant had in selecting, implementing, and evaluating tasks and any prior work the participants' students had done on reasoning-and-proving. The post-data collection interviews were driven by the data analysis. The research questions, grouped with the data that was collected, the coding system, and the analysis is shown in Table 3.3 (note: "RP" in Table 3.3 means "reasoning-and-proving").

Table 3.3 The Research Questions, Related Collected Data, Coding Systems, and Analysis

1.) To what extent did participants select reasoning-and-proving learning opportunities in the form of tasks?

Question	Data	Coding	Analysis
To what extent does the textbook include tasks that have the potential to engage students in reasoning-and-proving?	The tasks in the portion of the textbook that will be implemented during the data collection period	All tasks were coded using the Thompson, Senk and Johnson (2012) framework. Tasks that did not fit the framework were coded as non-RP. Appendix C	The percent of textbook tasks that had the potential to engage students in reasoning-and-proving were calculated. In addition, the percentage of textbook tasks in each category (Thompson, Senk and Johnson (2012) was calculated.
To what extent did the participant select tasks for instruction that had the potential to engage students in reasoning-and-proving?	All of the tasks that the teacher selected (used during or outside of class) during the data collection period Appendix F Appendix G	All tasks were coded using the Thompson, Senk and Johnson (2012) framework. Tasks that did not fit the framework were coded as non-RP. Appendix C	The percentage of RP tasks out of the selected tasks were calculated. In addition, the percentage of tasks used in each category (Thompson, Senk and Johnson (2012)) were calculated.
To what extent did the participant modify tasks to affect the tasks' potential to engage students in reasoning-and-proving?	Any task that was modified with respect to reasoning-and-proving from any source, used during the data collection period. Appendix F Appendix G	Modified tasks were coded for changes in reasoning-and-proving potential. Appendix H	The number of tasks that were modified was reported. In addition, the rationales provided by the participants regarding the modifications were reported.
What were the sources of the tasks that participants selected for instruction?	The Reasoning-and-Proving Task Log Sheet that lists the source of all reasoning-and-proving tasks used during the data collection period Appendix G	The tasks were coded as: (1) taken directly from the textbook; (2) modified from the textbook; (3) used in the CORP course: (4) taken or adapted from ancillary resources; or (5) created by the	The percents of tasks used in each category were calculated. In addition, the number of reasoning-and-proving tasks taken directly from the textbook were compared to the number of tasks in the textbook

		teacher	that have the potential to engage students in reasoning-and-proving.

2.) To what extent were participants able to implement reasoning-and-proving tasks?

Question	**Data**	**Coding**	**Analysis**
To what extent were participants able to maintain the level of cognitive demand of the reasoning-and-proving task during implementation?	Three classroom artifact packets of students' work on reasoning-and-proving tasks (one which asks students to show that something is "always true") and completed student work cover sheets for each of the three class-sets of work. Appendix I Appendix J	The three tasks were coded using the IQA rubrics for task potential and implementation. Appendix K	Each pair of codes for task potential and task implementation was reported. Average scores from each IQA rubric across the three tasks for each teacher were calculated.

3.) To what extent were participants able to accurate evaluate their students' reasoning-and-proving products?

Question	**Data**	**Coding**	**Analysis**
To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?	The completed classroom artifact packet cover sheets (which included the rubrics the teacher used to evaluate the students' work) for each of the three class sets of work. Appendix J	The rubrics submitted with the classroom artifact packets cover sheets were coded using the IQA rubric for clarity and detail of expectations. Appendix K	The quality (judged by clarity and detail of expectations) of the rubrics used by the teachers was reported.

To what extent did participants apply the core elements of proof in evaluating their students' reasoning-and-proving products?	The student work on reasoning-and-proving tasks which asked students to show that something is "always true" and completed classroom artifact packet cover sheets (which included the rubrics the teacher used to evaluate the students' work) for each of the three class-sets of work. Appendix J Appendix K	The students' reasoning-and-proving products were scored using the core elements for proof (which is listed in the IQA rubric for clarity and detail of expectations). Appendix K	A comparison was made between the teacher's and researcher's scores of the students' work.
In what ways did participants communicate expectations regarding what was required to produce a proof to students?	Background Interview question 3 and the classroom artifact packet cover sheet question 2. Appendix J Appendix N	The combination of the background interview questions and the classroom artifact packet cover sheet were coded using the IQA rubric for communication of expectations. Appendix K	The pattern of communication of expectations was examined (i.e. did the communication of expectations increase across the tasks?).

3.2.1 Classroom textbook or curriculum guide

The participants in the study were asked to select a second semester unit that included reasoning-and-proving activities. From the selected unit, the participants were asked to select a 15-day contiguous task collection period in consultation with the researcher. While the 15 day period was arbitrary, it should have provided a reasonable section of time in which to get an accurate picture of the instruction provided by each participant (assuming the modes of instruction fluctuate throughout a unit). This study examined four aspects of how participants selected reasoning-and-proving learning opportunities in the form of tasks: the availability of reasoning-and-proving tasks in the participants' textbook, the reasoning-and-proving tasks selected by the participant, the modifications the participant made to the selected tasks, and the source of the reasoning-and-proving tasks selected by the participant.

The Reasoning and Proof Standard in *Principles and Standards for School Mathematics* (NCTM, 2000) recommends that secondary students are given opportunities to make and investigate mathematical conjectures, develop and evaluate mathematical arguments and proofs, and select and use various types of reasoning and methods of proof. Any published curriculum that claims to incorporate the NCTM Standards (2000) should contain explicit evidence of these types of activities offered to students. The Standards also recommend that the curriculum and instruction offered to students helps them recognize reasoning-and-proof as a fundamental aspect of mathematics, but it is difficult to infer that recommendation from a textbook. Even with lesson plans, the level of inference regarding how the participant engaged students in discussions about properties would be too high to be accurate without observations. While Thompson, Senk and Johnson (2012) examined both the properties in the narrative portions and the exercises in the textbook chapters they studied, only the exercise portions of the textbook that were used during the collection period were examined in the study described herein.

This study assumed that each participant had been assigned or had selected a primary textbook to use in the classroom. The portion of the textbook that was used during the 15-day data collection period was identified by the participant. Any task that appeared in the exercise portions of the textbook—as opposed to the teacher's edition or the narrative portion of the student edition—was coded. In order to code these tasks, a copy of the student edition of the textbook was obtained by the primary researcher.

3.2.2 Task log sheets of relevant tasks

The tasks that the participants selected provided information about the extent to which the participants recognized and valued reasoning-and-proving. Before the start of the data-collection period, participants were asked to select a unit that involved reasoning-and-proving. During the 15-day data collection period, the participants filled out a task log sheet every day (Appendices F and G). In addition to listing every task used during class or assigned for homework, the participants listed the source of the task, identified whether or not the task was modified, determined if the task was a reasoning-and-proving task, recorded how much time was spent on the task, and recorded the purpose the task.

It was expected that the number of reasoning-and-proving tasks in the participants' textbooks was small (Lithner, 2004; G. J. Stylianides, 2009; Thompson, Senk & Johnson, 2012), which increased the likelihood of the participant having to modify tasks to increase (or create) the tasks' reasoning-and-proving potential during the data collection period. In order to examine how and why a participant modified a task, the participants were asked to provide the original task, the modified task, and their rationale for modifying the task.

It was also sometimes necessary for participants to search for additional sources for reasoning-and-proving tasks. Five sources of tasks were possible: taken directly from the textbook, modified from the textbook, used in the CORP course, taken or adapted from ancillary resources, and created by the participant. The data collected was the participants' identification of the sources of the tasks they selected for instruction as listed on the task log sheet each day during the data collection period. For tasks that were sourced beyond the textbook, the participant was to attach a copy of the task to the task log sheet for the relevant day.

3.2.3 Classroom artifact packets

The quality of instruction students receive in their classes is the most important school factor impacting student learning (Sanders & Horn, 1994). The directions and work the teacher assigns is a window into what instruction actually occurred in the classroom (Matsumura et al., 2006). Matsumura listed several reasons for collecting samples of teachers' assignments along with samples of student work: assignments offer information about an entire instructional cycle (communicate, practice/enact skills, feedback), they provide insight into the opportunities students have to produce individual work, and they provide further checks on the rigor of enacted tasks. These reasons frame the argument that collecting samples of teachers' assignments with corresponding work and evaluating the sets with the Instructional Quality Assessment (IQA) (Matsumura et al., 2006) provides a reasonable measure of instructional quality. The IQA is discussed in more detail in the data coding and analysis section.

The directions for the collection of these items were adapted from the directions for the collection of data from the IQA (see Appendix K). The adaptations were made to fit the needs of the current study (focus on reasoning-and-proving) and do not affect the validity of the instrument (M. Boston, personal communication, November 16, 2012). The classroom artifact packets contained a task, grading criteria recorded on the student work cover sheet, and 6 samples of student work (see Appendix I and J). At least one of the tasks was to have required students to construct a proof (i.e., show that something was *"always true"*). From each set of student work, the participants were asked to identify six samples: two samples that exceeded expectations, two samples that met expectations, and two samples that failed expectations.

The participants were asked to complete a cover sheet for each classroom artifact packet (see Appendix J). The cover sheet questions provided information about the nature of that task (typical or not), the directions and expectations for the task work shared with the students, implementation, and criteria for evaluation. Thus, the data collected from the classroom artifact packets informed the extent to which teachers were able to maintain the level of cognitive demand of the reasoning-and-proving tasks during implementation and the extent to which the participants were able to accurately evaluate their students' reasoning-and-proving products from those tasks.

3.2.4 Interview Questions

The background interview questions fell into three categories: school descriptive information, freedom to choose curriculum, and prior reasoning-and-proving work with students (Appendix N). The school descriptive information provided information about the size of the school, teacher-to-pupil ratio, setting, and which mathematics class was chosen for this study. Such information helped identify the challenges that each participant faced in engaging their students in reasoning-and-proving activities. Data on the freedom to choose curriculum data provided insight into how the participants in the current study approached incorporating reasoning-and-proving into their first- and second-year classrooms. No teacher acts completely independently in a school system; not only is a teacher's practice influenced by his own beliefs, goals, content and pedagogical content knowledge (Schoenfeld, 1998), but it is also influenced by collaboration with other teachers (Sadri, 2008), students (Raudenbush, Rowan, & Cheong, 1993), curriculum (Remillard, 2000), and the leadership structure of the school (Sather, 1994). If a participant was surrounded by supportive teachers who also wish to incorporate reasoning-and-proving into their classes or who have already engaged students in reasoning-and-proving in a course prior to the participants' course, the extent to which the participants selected, implemented and evaluated reasoning-and-proving tasks would probably be different than a participant who was isolated or surrounded by unsupportive students.

In order to examine how the participants shared their expectations for quality reasoning-and-proving work with their students, information on reasoning-and-proving activities on which students worked prior to the data collection period was gathered. Ideally, the participants would have already helped students develop class criteria for proof that contained the core elements and that criteria would be readily available to students for reference. If students worked on a succession of reasoning-and-proving tasks throughout the year, there should be evidence of the students' deepening understanding of reasoning-and-proof. The third set of background interview questions coupled with the classroom artifact packets informed the pattern of how the teacher communicated her expectations throughout the data collection period (and possibly the course).

Follow-up interviews were done after the initial analysis of the data contained on the task log sheets and classroom artifact packets. The questions were designed to further probe the participants' thinking and reflections of the reasoning-and-proving activities and to clarify any questions that arose from the analysis. Each interview also asked participants how they would define proof to a colleague and to a student.

The four sources of data—textbooks, task log sheets with relevant tasks, classroom artifact packets, and interview questions—provided enough information to determine the extent to which the participants are able to select, implement, and evaluate student products of reasoning-and-proving tasks. In the next section, the coding systems and analysis are described for each source of data and how the collection will inform each research question.

3.3 DATA CODING AND ANALYSIS

Four coding systems were used to analyze the data in the current study. The reasoning-and-proving task codes used by Thompson, Senk and Johnson (2012) were used to determine whether or not a task is a reasoning-and-proving task. The IQA rubrics were used to assess the potential and implementation of tasks, the clarity and detail of expectations of the participants, and the communication of those expectations to students. The other coding systems were for determining the changes in reasoning-and-proving potential of modified tasks and a simple system for categorizing the sources of tasks that the participants selected.

3.3.1 Selecting reasoning-and-proving tasks (Research Question 1)

Two sources of data—the classroom textbooks and the task log sheets with relevant tasks—were used to answer the four sub-questions of the research question, "to what extent did participants select reasoning-and-proving learning opportunities in the form of tasks." The data were from the 15 days of instruction selected by the participant. The coding and analysis were identical for determining the results of the availability of reasoning-and-proving tasks in the classroom textbook and of the selection of reasoning-and-proving tasks by the participant to be used for instruction. For the task modifications and the sources of tasks, only the task log sheets of relevant tasks were used for coding and analysis.

When Thompson, Senk, and Johnson (2012) conducted their study of reasoning-and-proving learning opportunities in textbooks, they chose to use *exercises* as their unit of analysis instead of *tasks.* Stein, Grover, & Hanningsen (1996) defined a *mathematical task* as

> a classroom activity, the purpose of which is to focus the students' attention on a particular mathematical idea. An activity is not classified as a different or new task unless the underlying mathematical idea toward which the activity is oriented changes. Thus, a lesson is typically divided into two, three, or four tasks rather than into many more tasks of shorter duration (p. 460).

Because Thompson, Senk, and Johnson (2012) intended to compare the amount of available reasoning-and-proving available in textbooks from different publishers and each publisher labeled sets of exercises in different ways, the researchers counted the "total number of exercises when all the labeled parts were counted separately" (p. 264). In other words, if a textbook practice problem had parts a-c, Thompson, Senk, and Johnson (2012) counted this problem as three *exercises* in order to standardize the different formats used by different publishers. Since the study described herein compares the amount and types of available reasoning-and-proving opportunities in textbooks to the amount and types of selected reasoning-and-proving opportunities, this study will use *exercises* for the coding and analysis used to answer the first research question for the remainder of this document. From this point forward, the first research question will be "to what extent did participants selected reasoning-and-proving learning opportunities in the form of exercises."

3.3.1.1 Available and selected exercises

Both the exercises from the practice portions of the textbook and the selected exercises as listed on the task log sheet will be coded using Thompson, Senk & Johnson's (2012) categories. The category codes relate to the task action required on the student's part: make a conjecture, investigate a conjecture, develop an argument, evaluate a given argument, find a counterexample, determine the error in reasoning in a given solution, and create the outline of an argument. These categories map onto G. Stylianides' mathematical components of his analytical framework of reasoning-and-proving (see Appendix C), and examples of each type of action can be found in Appendix D. Any task that does not fit into one of these categories will be coded as "non-reasoning-and-proving" (non RP).

The first analyses informed whether the participant capitalized on all of the available reasoning-and-proving learning opportunities in the textbook. The analysis of this coded data took the form of four calculations. To answer the sub-question "to what extent does the textbook include exercises that have the potential to engage students in reasoning-and-proving," the overall percentage of textbook exercises that have the potential to engage students in reasoning-and-proving was calculated:

$$\frac{\#\ RP\ textbook\ exercises}{\#\ textbook\ exercises} * 100$$

To determine the types of reasoning-and-proving activities available to students in these selected sections of the textbook, the percentage of textbook exercises in each of the categories make a conjecture, investigate a conjecture, develop an argument, evaluate a given argument, find a counterexample, determine the error in reasoning in a given solution, or create the outline of an argument was also calculated. For instance, to determine the percentage of reasoning-and-proving tasks in the textbook that directed students to develop a general argument, the following expression was used:

$$\frac{\#\ DG\ exercises\ in\ textbook}{\#\ of\ RP\ exercises\ in\ textbook} * 100$$

Similar calculations were used to answer the second sub-question, "to what extent did the participants select exercises for instruction that have the potential to engage students in reasoning-and-proving?" For this calculation, however, the percent of reasoning-and-proving exercises out of the *selected exercises used in instruction* was found instead of the available textbook exercises:

$$\frac{\#\ RP\ exercises\ selected}{\#\ selected\ exercises} * 100$$

If the percentage of reasoning-and-proving tasks available in the textbook was lower than the percentage of reasoning-and-proving exercises selected by the participant, then the participant increased the potential of reasoning-and-proving learning opportunities to students.

The types of selected reasoning-and-proving exercises were examined by calculating the percentage of selected exercises in each category out of the selected reasoning-and-proving exercises. For example, the percentage of selected exercises that asked students to develop a general argument out of the total number of selected reasoning-and-proving exercises was found with the following expression:

$$\frac{\#\ DG\ exercises\ selected}{\#\ of\ RP\ exercises\ selected} * 100$$

Ideally, the participants selected an assortment of reasoning-and-proving types of tasks, providing students with well-rounded learning experiences for reasoning-and-proving. On the task log sheet, the participants were asked to record the time spent on the exercise and the purpose of the exercise. If reasoning-and-proving activities commanded the lion's share of the instructional time (both in and out of class), it was assumed that the participant highly valued reasoning-and-proving activities. If very little time was spent on reasoning-and-proving, especially during a 15-day cycle in which the participant claimed to be working on reasoning-and-proving, it suggests that the participant did not value reasoning-and-proving, was not prepared to select and implement reasoning-and-proving tasks, or had no control over the daily curriculum.

3.3.1.2 Changes in the reasoning-and-proving potential of exercises

During the CORP course, the participants studied several sets of tasks (original and modified versions) and created a list of strategies they could use to make their own modifications to tasks. One type of modification involved stripping away unnecessary scaffolding from a task and instead asking students to investigate the task's concept by generating some examples, making observations, and making their own conjectures. For example, a teacher could remove the structure of a fill-in-the-blank geometry proof template and ask seasoned geometry students to practice writing proofs for conjectures in the format of their choosing. Another type of modification involved changing the verbs in the prompts from only making generalizations (identifying a pattern, make a conjecture) to making an argument (non-proof or proof). Participants were reminded to modify only the task's directions and not the task's mathematics. In other words, take a task that asks students to make a conjecture about parallel lines and a transversal and modify the direction to include "show that your conjecture is always true" but do not change the mathematics of the task to focus only on alternate interior angles.

While the modification work in the fourth chapter of the CORP course was intended to help participants *increase* the reasoning-and-proving learning opportunities of tasks, it is possible to modify a task to *decrease* the reasoning-and-proving potential. According to G. J. Stylianides (2010), there is a hierarchy of reasoning-and-proving activities. From low to high they are: identify a pattern, make a conjecture, provide a rationale, and construct a proof (see Figure 2.1). For the study reported herein, any transition from a lower level to a higher level was considered an increase in reasoning-and-proving potential. For example, a geometry textbook task might direct students to investigate the conjecture, "the sum of the angles in a triangle is 180 degrees." If the participant modified the task to direct the students to "investigate the conjecture…and show that the claim is *always true*," then the participant increased the level of reasoning-and-proving of the task. On the other hand, if a participant watered down a proof task by asking students to stop at a conjecture, then the participant modified the task to lower the reasoning-and-proving potential of the task. Any exercise that a participant modified was self-identified on the task log sheet (see Appendix G). The changes in exercises were identified by comparing the original and modified versions of the exercises as well as considering the rationale for the modification (see Appendix F).

A teacher can have many reasons for modifying exercises. The teacher may wish to increase the reasoning-and-proving potential of the exercise, or decrease the number of representations of an exercise in order to focus on a particular skill. Not all modifications involve reasoning-and-proving potential, and not all of the exercises selected by the participant in this data collection period were reasoning-and-proving exercises. The only exercises that were counted and analyzed were exercises that either started as reasoning-and-proving exercises (as assessed by the raters) and were modified, or exercises that were not originally reasoning-and-proving and were modified to include reasoning-and-proving (as assessed by the raters). As such, four modification possibilities exist. First, an exercise that had the potential to engage students in reasoning-and-proving could be modified to *lower* the reasoning-and-proving potential of the exercise. These exercises were assigned the code -1 (see Appendix H). An exercise that had the potential to engage students in reasoning-and-proving could modified, but not in a way that affects the reasoning-and-proving potential of the exercise. For instance, a participant may change the wording of the problem to ask students to record specific information, such as formulas or definitions. A reasoning-and-proving exercise that was modified in such a way that the modification had a neutral effect on the reasoning-and-proving potential of the exercise was assigned the code 0. Exercises that had the potential to engage students in reasoning-and-proving that were modified to *increase* the reasoning-and-proving potential were assigned the code +1. Finally, it is possible that an exercise originally did not have the potential to engage students in reasoning-and-proving but the participant modified the task to provide such potential. It was the researcher's assumption that transforming a non-reasoning-and-proving task into a reasoning-and-proving task is more complicated and takes more skill than merely changing the directions of a task already primed for reasoning-and-proving; as such, these transformed tasks were assigned the code +2. If a mixture of all of these types of modifications was made by a participant, the sum of the modification codes was reported. This sum indicated whether the participant generally increased or decreased the reasoning-and-proving potential of exercises through modifications.

The researcher did not expect the participants to modify many of the tasks, therefore, the ways exercises were modified was described in Chapter 4. In many cases, the original and the modified versions of exercises were also provided in Chapter 4. This evidence, coupled with the modification code scores, informed the extent to which the participants modified exercises to affect the exercises' potential to engage students in reasoning-and-proving.

3.3.1.3 Sources of Exercises

On the task log sheet (Appendix G), participants were asked to record the source of the tasks they selected for instruction. The sources of the tasks coded in one of the following ways: taken directly from the textbook, modified from the textbook, used in the CORP course, taken or adapted from ancillary resources (such as the internet or a fellow teacher), or created by the participant. The percentage of tasks used from each source was then calculated:

$$\frac{\#\ unmodified\ exercises\ selected\ from\ textbook}{\#\ of\ selected\ exercises} * 100$$

$$\frac{\#\ modified\ exercises\ selected\ from\ textbook}{\#\ of\ selected\ exercisese} * 100$$

$$\frac{\#\ exercises\ selected\ from\ CORP\ course}{\#\ of\ selected\ exercises} * 100$$

$$\frac{\#\ exercises\ selected\ from\ ancillary\ resources}{\#\ of\ selected\ exercises} * 100$$

$$\frac{\#\ modified\ exercises\ created\ by\ the\ teacher}{\#\ of\ selected\ exercises} * 100$$

The extent to which participants are able to find exercises in their own curriculum materials, mine outside sources (including fellow teachers), or create exercises themselves spoke to the extent to which the participants were able to offer reasoning-and-proving learning opportunities to their students.

3.3.2 Implementing reasoning-and-proving tasks (Research Question 2)

The quality of instruction students receive in their classes is the most important school factor impacting student learning (Sanders & Horn, 1994). Cognitively rigorous activities offer students potentially challenging tasks that require complex and non-algorithmic thinking, engage students in creating mathematical meanings, are implemented at a similarly high level, and require responses that explain the validity of strategies. Reasoning-and-proving tasks, as defined by G. J. Stylianides (2010), are cognitively rigorous. The data used to inform the question, "to what extent were teachers able to maintain the level of cognitive demand of the reasoning-and-proving task during implementation?" were the three classroom artifact packets, each containing a task, six samples of student work (2 labeled exceeded expectations, 2 labeled met expectations, and 2 labeled failed expectations), and the student work cover sheet. Each packet was selected, gathered, and generated by the participant.

In order to determine if a participant maintained the level of cognitive demand of a task during implementation, both the original level and the implemented level of cognitive demand of the task had to be determined. Since some of the IQA Mathematics Toolkit rubrics were designed to measure instructional quality without observations but instead with classroom artifact packets, the classroom artifact packet rubrics were used as a tool for measuring level of cognitive demand. According to Matsumura and her colleagues, the directions and work the teacher assigns is a window into what instruction actually occurred in the classroom (Matsumura et al., 2006). Specifically, the IQA Mathematics Toolkit looks at the cognitive rigor in tasks, the rigor in the implementation of the tasks, the rigor in the student responses to the tasks, and the clear and detailed expectations that have been communicated to students (See Appendix K). Since reasoning-and-proving activities are cognitively demanding, the results of the IQA rubrics informed the extent to which participants are able to implement reasoning-and-proving tasks. The IQA rubrics were applied to the three tasks, cover sheets, and accompanying student work from the classroom artifact packets.

The first IQA rubric—task potential—helped determine if the tasks selected by the participants had the potential to engage students in reasoning-and-proving (see Appendix K). No participant selected a task for their classroom artifact packets that was not a reasoning-and-proving task. According to the task potential rubric, a reasoning-and-proving task that has the highest potential to engage students in cognitively demanding work cannot be predictable and directs the students to do some or all of the following: identify patterns and form generalizations based on those patterns, make or investigate conjectures and support conclusions with mathematical evidence and/or create a proof or find a counterexample, and evaluate an argument or explain how to outline an argument of a particular type.

A task that directed students to identify patterns but not to form generalizations or directed students to make conjectures but not support conclusions with mathematical evidence or create proofs received a score of 3 on the task potential rubric. A code of 3 was also be used for tasks that were well beyond the conceptual reach of the students (e.g., required techniques or concepts not yet learned, such as proof by contradiction for an Algebra 1 student). A code of 2 was assigned to formulaic tasks that were unambiguous and focused on producing correct answers rather than mathematical understanding. Any task that did not require students to make connections or develop meanings (e.g., reproduce memorized facts) received a code of 1. A sample of tasks for each of the codes in the task potential rubric can be found in Appendix L.

The second IQA rubric—task implementation—closely mirrors the rubric for task potential. The codes for this rubric inform whether the students were held to the cognitive demand of the task and then indicate the level of implementation. In other words, if a task directed students to make a conjecture and support their conclusions with a proof or counterexample, did the students' work show evidence that the students made conjectures and supported their conclusions? If so, then the participant implemented the reasoning-and-proving in a way that maintained the task's original level of cognitive demand. Examples of each level of task implementation can be found in Appendix M. Information about the task implementation was requested on the classroom artifact packet cover sheet (Appendix J) in the form of two questions: was the task implemented differently than planned, and what was the implementation successful? The answers to these questions provided additional information regarding the participants' implementation of the reasoning-and-proving tasks.

In qualitative research, reliability is defined as a fit between what the researcher recorded as data and what actually occurred in the setting under study, rather than the literal consistency across different observations (Borgen & Biklen, 2003). The IQA has already met this standard of reliability because the data used will be selected and sent by the participants themselves and as such accurately reflect their practice (Clare & Aschbacher, 2001; Matsumura, Garnier, Pascal, & Valdés, 2002). To ensure reliability in coding, two raters who were familiar with both the IQA and reasoning-and-proving applied the IQA rubrics to every classroom artifact packet. Any discrepancies were resolved by consensus, and additional information was gathered from the participants when necessary.

For each task in the classroom artifact packet, the pair of task potential and task implementation codes was reported, which revealed trends in maintaining or declining cognitive demands. This provided an overall picture as to the participants' extent to which they were able to implement reasoning-and-proving tasks.

3.3.3 Evaluating students' reasoning-and-proving products (Research Question 3)

3.3.3.1 Core elements of proof

The core elements of proof are: the proof's argument must show that the conjecture or claim is (or is not) true for *all* cases, the statements and definitions that are used in the argument must be ones that are true and accepted by the community (because they have been previously justified), and the conclusion that is reached from the set of statements must follow logically from the argument made. The following aspects of proof are *not* core elements: type of proof, form of the proof, the representation used, and the explanatory power of the proof. If the participants internalized the core elements of proof during the CORP course, one would expect the core elements to be the largest factor in the participants' criteria for judging the validity of a proof and that the participants applied those core elements of proof in evaluating their students' reasoning-and-proving products. In order to determine the extent to which this happened, the classroom artifact packets were examined. The classroom artifact cover sheet (see Appendix J) included directions for the participants to submit their scoring rubric for the student work, and the classroom artifact packets were each to contain six pieces of student work: two evaluated as exceeding expectations, two evaluated as meeting expectations, and two evaluated as failing expectations. Participants were also asked how they would define proof to a colleague and to a student. These definitions were reported verbatim and were analyzed for the core elements of proof. Any additional criteria (non-essential) was also reported.

According to Boston and Wolf (2006), instructional quality depends on the academic rigor in a teacher's expectations and the clarity, detail, and communications of those expectations. As such, the evaluation criteria (rubrics) that the participants used to evaluate their students' work were coded using the fifth IQA rubric: clarity and detail of expectations (see Appendix K). An evaluation system that contained clear and elaborated expectations for the quality of student work was coded as 4. For proof tasks, the expectations must include the core elements of proof for a 4. Expectations that were less clear or treated non-core elements of proof as essential (i.e. format of the proof) received a code of 3. Unelaborated expectations or no expectations specified in the evaluation system received a code of 2 or 1, respectively. The results of the rubrics were reported and examined for trends regarding the extent to which the participants' criteria for judging the validity of their students' reasoning-and-proving products contained the core elements of proof.

To determine the extent to which the participants applied the core elements of proof in evaluating their students' reasoning-and-proving products, only tasks from the artifact packets which asked students to prove (or “show that this is always true”) were coded in this section. The six pieces of student work from each classroom artifact packet that the participant labeled “exceeded expectations”, “met expectations”, and “failed expectations” were also scored by the two raters (the primary researcher and a trained graduate student). The raters scored each sample of student work for each proof task according to the core elements of proof. Both the participant’s and the raters’ assessment of each student proof was reported. The comparison of the researcher’s rating (based on the core elements of proof and the participant’s rating revealed the extent to which the participants applied the core elements of proof in evaluating their students’ reasoning-and-proving products.

3.3.3.2 Communication of Expectations

Communicating expectations to students is a mark of instructional quality (Boston & Wolf, 2006). As such, the participants’ classroom artifact packet cover sheets and the background interview questions were mined in order to determine if and how participants developed or shared the criteria for proof with their students (see Appendices J and N). Evidence of developing or sharing would include posters of proof criteria, reference sheets in the students’ notebooks, or a description of an actual lesson in which students developed criteria for proof based on the core elements of proof. Such evidence was requested in the participants’ background interview questions and in the student work cover sheet (“please explain any expectations you relayed to your class”).

The IQA rubric for communication of expectations was based on when the participant shared the criteria and whether or not high-quality work was modeled. Participants who discussed their expectations (i.e. criteria for proof) and modeled high-quality work in advance of the task received a communication code of 4. Participants who discussed the criteria but provided no models were assigned a 3, and participants who presented the criteria to their students without discussion were assigned a 2. Participants who did not share their expectations at all or shared them after the task was completed were assigned a 1.

This analysis provided insight into the teachers' sense of the potential of recognized and selected tasks, how the tasks were implemented, and the standards of rigor to which the teachers held the students accountable. The three classroom artifact packets allowed for the assessment of the academic rigor, clarity, detail, and communication of the teachers' expectations for their students regarding proof tasks. The resulting detailed picture of the participants' enactment of reasoning-and-proving in their classrooms allowed for the observation of trends across participants with respect to what they internalized from the CORP course (following Yin's (1993) replication logic).

The findings for each participant was described as separate cases. Participants taught different courses in different contexts and it was assumed that they internalized the CORP course information in different ways. Describing the extent to which each participant selected, modified, sourced, and evaluated the products of tasks in separate cases was appropriate. Once a rich description of each participant has been provided, the trends across participants were described.

3.4 VALIDITY AND GENERALIZABILITY

The purpose of this study was not to evaluate the effectiveness of the CORP course or to assess the effectiveness of particular beginning teachers (Bogdan & Biklen, 2003; Merriam, 1998). However, I do hope that this study contributes to our understanding of how early-career teachers who experienced targeted professional development enact reasoning-and-proving tasks in their classrooms. To ensure the validity of the conclusions drawn in this study, standards for collecting tasks and student work were adopted from peer-reviewed research (Clare & Aschbacher, 2001; Matsumura et al., 2002) and the major coding systems were also from peer-reviewed research (G. J. Stylianides, 2010; Thompson, Senk & Johnson, 2012).

While this study is not designed to be generalized to all teachers in all districts, the fact that the participants taught in different types of schools in two different states suggests that any trends across all cases can be transferred to new settings (Stake, 1978 cited in Schofield, 2002; Lincoln and Guba, 1985). Although my study contained only four participants, the limited number of participants allowed me to provide rich detail about how the participants enacted reasoning-and-proving in their classrooms.

Two raters coded each of the exercises in the textbook and on the participants' task log sheets with respect to reasoning-and-proving potential and then type of reasoning-and-proving. The incident of discrepancies was small, and most discrepancies were resolved through consensus. 2,295 textbook exercises were coded; there was agreement on 2,285 exercises (99.6% agreement). Of the ten tasks on which there was no agreement, seven were exercises that one rater judged to have the potential to engage students in reasoning-and-proving while the other three were disagreements as to the type of reasoning-and-proving (e.g., make a conjecture or the combination of make a conjecture and develop an argument). For the 1,453 exercises selected by participants, the few discrepancies were resolved by consensus (100% agreement). Thus, overall, there was a 99.7% agreement on the exercises available in the textbook or selected by the participant. There was 100% agreement on the IQA rubric scores, once discrepancies were resolved by consensus.

4.0 DATA AND ANALYSIS

This chapter shares the data and analysis which was described in Chapter Three and was used to answer the three research questions of this study. The data and analysis is organized in four case studies, one for each participant. The first two case studies are Karen and Uma, both second-year educators teaching geometry in urban districts, the former with a reform curriculum and the latter with a traditional curriculum. The second two case studies are Sidney and Jonathan, both first-year educators teaching Algebra 1, the former in an affluent suburban district and the latter in a start-up charter school with urban students. These cases provide details about how these participants selected, implemented, and evaluated the student work products from reasoning-and-proving exercises. The final section of the chapter pulls data together from all four cases and analyzes the data for trends.

Throughout this chapter, the word "exercise" denotes a single activity. Thompson, Senk, and Johnson (2012) used this method to standardize the different ways textbook publishers label practice problems. Thus, a problem with subparts (e.g., #4a, 4b, 4c, and 4d) was counted as four separate exercises; each part was evaluated on its reasoning-and-proving potential. The exercises selected by the participants were also counted in this way.

As a reminder, this study analyzes the opportunities teachers who participated in a university course focused on reasoning-and-proving gave their students to engage in reasoning-and-proving. The course included opportunities to learn about the importance of reasoning-and-proving in mathematics, use frameworks and criteria to examine and evaluate student work, and consider how to implement reasoning-and-proving tasks. This study sought to learn how the teachers used knowledge gained in the course in the context of their classrooms. In particular, this study examines the following questions:

1.) To what extent did participants select reasoning-and-proving learning opportunities in the form of exercises?
    a. To what extent does the textbook include exercises that have the potential to engage students in reasoning-and-proving?
    b. To what extent did the participants select exercises for instruction that have the potential to engage students in reasoning-and-proving?
    c. To what extent did the participants modify exercises to affect the exercises' potential to engage students in reasoning-and-proving?
    d. What were the sources of the exercises participants selected for instruction?
2.) To what extent were participants able to maintain the level of cognitive demand of the reasoning-and-proving tasks during implementation?
3.) To what extent were participants able to accurately evaluate their students' reasoning-and-proving products?
    a. To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?
    b. To what extent did participants apply the core elements of proof in evaluating their students' reasoning-and-proving products?

c. In what ways did participants communicate expectations regarding what is required to produce a proof to students?

The case studies and trend section present and analyze data for each of these questions, in the order that the questions are listed in this paragraph.

4.1 KAREN

Karen was a second-year teacher at a science and technology magnet school in a large urban district in Pennsylvania when this study was conducted. The school is fairly new and includes grades 6-12. There are about 60 students in a graduating class; the school has a capacity for about 100 students per grade in 9^{th}-12^{th} grades. Karen teaches both a semester long geometry class for average students and two full-year geometry classes for struggling students. Each class period is 80 minutes long except on Wednesdays, when students lose 15 minutes per class in order to make professional development time available to teachers.

Karen selected her second semester geometry class for this study. According to Karen, her semester students have better math skills, work ethic, and behavior than her full-year geometry students, and are willing to come in for extra help when necessary. The curriculum used by Karen's district is the Center for Mathematics Education (CME) Project which was created by the Education Development Center, Inc. with partial funding by the National Science Foundation. This is the second year that the district (and Karen) has used CME's *Geometry* (2009); the district also adopted CME *Algebra* this past year. Karen reported that her district's focus on the Common Core standards, CME's focus on mathematical understanding, and the collaborative nature of her department all resonate with her university training (personal communication, February 20, 2013).

4.1.1 Selecting exercises (*RQ1: To what extent did participants select reasoning-and-proving learning opportunities in the form of exercises?)*

4.1.1.1 Available in textbook (*RQ1a: To what extent does the textbook include exercises that have the potential to engage students in reasoning-and-proving?)*

Karen selected part of CME's *Geometry: Congruence and Proof* unit for her 15-day data collection. Unlike a traditional textbook, Karen's reform textbook did not structure lessons with a distinct content description and explanation section followed by a distinct practice problem section, therefore, all exercises were counted regardless of intent. Also unlike most curricula, CME's *Geometry: Congruence and Proof* is full of reasoning-and-proving tasks. Out of the 436 textbook exercises available to Karen during her 15-day data collection period, 329 (75.5%) of them were coded as reasoning-and-proving. The tasks spanned all of the types of reasoning-and-proving listed by Thompson, Senk, and Johnson (212), with concentrations in Investigate a Conjecture and Develop an Argument (see Table 4.1). Depending on the directions of the exercises, some of the exercises were coded as more than one type (e.g., make a conjecture *and* develop an argument); therefore, the numbers in Table 4.1 is greater than 436. All of the textbooks and task log sheets of each participant contained exercises that were double- and triple-coded.

Table 4.1 Available Types of Reasoning-and-proving Exercises in Karen's Textbook

Type of RP Exercise	Available in Textbook*
Make a Conjecture	56 (12.8%)
Investigate a Conjecture	141 (32.3%)
Evaluate an Argument	4 (0.9%)
Correct a Mistake	4 (0.9%)
Develop an Argument	150 (34.4%)
Counterexample	14 (3.2%)
Principles of Proof	24 (5.5%)
Non-Reasoning-and-Proving	107 (24.5%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

Karen's unit was the only unit of any participant in the study to contain "Principles of Proof" exercises. For example, students were asked "why is poof so important in mathematics?" (CME *Geometry,* Unit 2, p. 112, #5) and "Draw a square and divide it into four equal parts. Write an argument that convinces a classmate that each part has the same area. Share your argument with a classmate. Is your argument convincing?" (CME *Geometry,* Unit 2, p. 115, #1). The latter exercise was coded as both Develop an Argument for "write an argument" and Principles of Proof for "is your argument convincing?"

4.1.1.2 Selected by Teacher (*RQ1b: To what extent did the participant select exercises for instruction that had the potential to engage students in reasoning-and-proving?*)

Karen did not assign every exercise available in her textbook, but she did create or find additional exercises to select; in total, she offered more reasoning-and-proving exercises to her students than were available in her textbook. Out of the 335 exercises assigned to her students, 256 (76.4%) of them were reasoning-and-proving and 79 were not, as rated by the coders. The exercises were largely from her curriculum and included explorations, practice exercises, maintaining skills exercises, mathematical reflections, discussion questions, exit slips to check understanding, chapter reviews, and chapter tests.

Of the 256 reasoning-and-proving exercises she selected, Karen labeled 243 of them as reasoning-and-proving and did not label the remaining 12 as reasoning-and-proving. Of the latter 12 exercises, 4 asked students to make a conjecture and 8 were Principles of Proof exercises. The Principles of Proof exercises were from Day 6 of Karen's data collection period (Problem 2.10: What Does a Proof Look Like?). For example, the exercise "why might it be important to have several different styles of writing proofs?" (Karen's Lessons Day 6, Wrap-up question 3) is a Principles of Proof exercise. Of the 79 exercises rated as non-reasoning-and-proving by the coders, Karen labeled 2 exercises as not reasoning-and-proving but labeled the remaining 77 as reasoning-and-proving. This data is reported with some caution, however, because on her Task Log Sheets, Karen sometimes grouped collections of exercises together and labeled the entire set as "reasoning-and-proving." For instance, Karen labeled the entire unit review (Day 14) and chapter exam (Day 15) as reasoning-and-proving, but the coders only rated some of the exercises in these collections as reasoning-and-proving. Close examination of Karen's labels on smaller sets of exercises does reveal some consistent patterns, however.

Exercises that asked students to find multiple angles in a figure (by applying properties such as the Alternate Interior Angle Theorem for parallel lines), record or apply definitions, identify the hypothesis and conclusion of a statement, or provide an explanation were labeled as reasoning-and-proving by Karen. While these skills are useful in reasoning-and-proving, alone they do not qualify as the type of reasoning that can lead to proof as identified by Thompson, Senk, and Johnson (2012). Figure 4.1 shows one example of an exercise that requires students to explain why four figures are not quadrilaterals; certainly knowing and being able to use definitions are part of the reasoning-and-proving process but *only* applying a definition is not sufficient to engage students in a reasoning-and-proving activity.

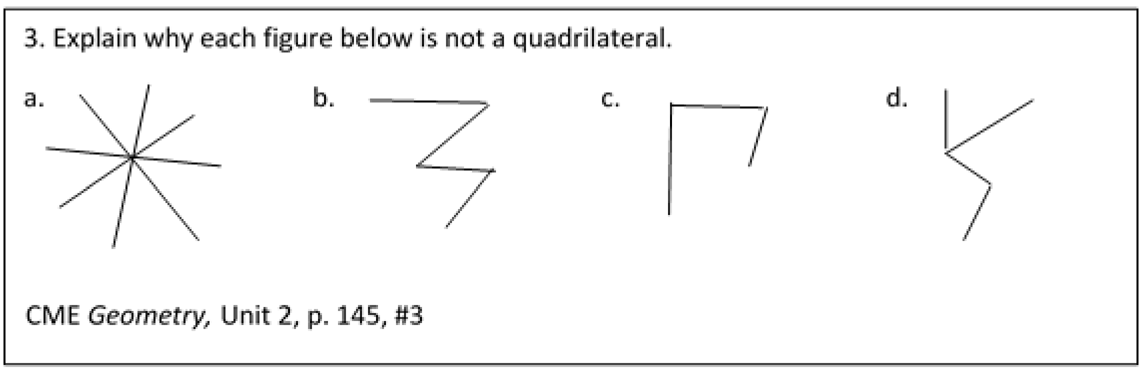


Figure 4.1 A non-reasoning-and-proving exercise mislabeled by Karen.

Table 4.2 lists the types of reasoning-and-proving exercises in which Karen engaged her students. Similar to what was available in the textbook, Karen offered a range of types with a concentration on Investigate a Conjecture (37.9%) and Develop an Argument (37.3%). These percentages are a little higher than the types available in Karen's textbook (32.3% and 34.4%, respectively). Figure 4.2 represents this comparison graphically.

Table 4.2 Selected types of reasoning-and-proving exercises by Karen.

Type of RP Exercise	Selected by Teacher*
Make a Conjecture	24 (7.2 %)
Investigate a Conjecture	127 (37.9 %)
Evaluate an Argument	1 (0.3%)
Correct a Mistake	2 (0.6 %)
Develop an Argument	125 (37.3 %)
Counterexample	16 (4.8 %)
Principles of Proof	18 (5.4 %)
Non-Reasoning-and-Proving	79 (23.6 %)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

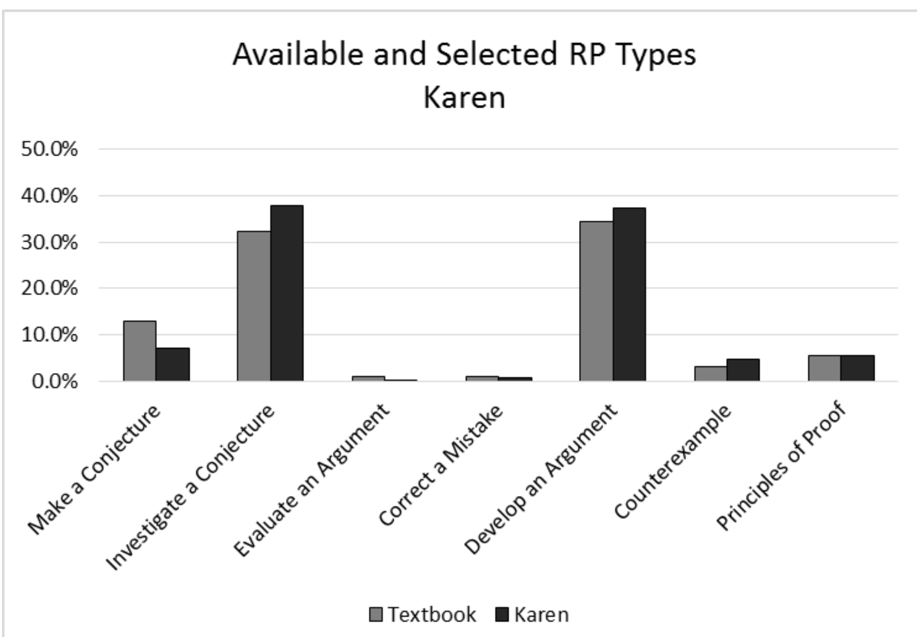


Figure 4.2 Comparison of available and selected types of reasoning-and-proving exercises for Karen.

4.1.1.3 Exercise modifications *(RQ1c: To what extent did the participant modify exercises to affect the exercises' potential to engage students in reasoning-and-proving?)*

Karen modified 29 exercises during her data collection period. The modifications were mainly to direct her students' mathematical behavior rather than change the exercises' reasoning-and-proving potential; thus, the modifications had a neutral effect on students' opportunities to engage in reasoning and proving (see Table 4.3). The types of exercise modifications made by Karen were:

- Karen constructed figures in GeoGebra that students used to lead 5 class explorations
- Karen added scaffolding to 3 exercises to help students think through the reverse proof process
- Karen directed her students to draw pictures and explain 21 Investigate a Claim-type exercises (fill-in-the-blank with *always, sometimes,* or *never*; see Figure 4.3)

Table 4.3 Frequency of Exercise Modifications Made by Karen

Original Exercise (any source)	Exercise as assigned by teacher	Code	Frequency
Reasoning-and-Proving Exercise	Exercise assigned, modified to LOWER RP	-1	0
Reasoning-and-Proving Exercise	Exercise assigned, neutral effect of modification	0	29
Reasoning-and-Proving Exercise	Exercise assigned, modified to INCREASE RP	+1	0
Non-Reasoning-and-Proving Exercise	Exercise assigned, modified to INCLUDE RP	+2	0

Original Exercise:

For Exercises 9-28, complete each sentence with *always, sometimes,* or *never* to make the statement true.

9. A parallelogram __?__ has two congruent sides.

17. A quadrilateral with one right angle is __?__a parallelogram.

(CME's *Geometry,* Unit 2, p.154 #9 and p. 155 #17)

Karen's Modification:

Draw pictures and explain for full credit.

Figure 4.3 Original and modified exercise from Karen's Lesson Day 12.

Karen did not lower the reasoning-and-proving potential of any exercise. In fairness, over three-quarters of the exercises selected by Karen were already reasoning-and-proving exercises. One could argue that no modifications were necessary to increase the amount of reasoning-an-proving that was available in the curriculum.

4.1.1.4 Exercise sources *(RQ1d: What were the sources of the exercises that participants selected for instruction?)*

Of the 335 exercises assigned by Karen, 77.3% were taken directly from the CME *Geometry* Unit 2. This was no surprise given the focus of this unit and the district curriculum guide provided to Karen. The district curriculum guide details the goals of each investigation (subunits of the CME curriculum), lists overarching questions, objectives, Common Core standards, provides assessing and advancing questions, and suggests assignments. All of these are closely aligned to the CME curriculum, which was tailored by the publisher for Karen's large urban district. The district curriculum guide is written for a year-long course; Karen had to modify it to fit her semester course, but her decisions were largely based on time constraints (the semester course did not have the same number of instructional hours as full-year, non-blocked courses taught in other buildings in the district) and a concern that her students needed more skill practice than the curriculum provides (personal communication, February 20, 2013).

When Karen did take or adapt exercises from ancillary sources, she used district exams from previous years or the RegentsPrep.org website, a repository of tasks designed to help New York secondary students prepare for their state assessment exams. Table 4.4 details the number of exercises (out of 335) from each type of source selected by Karen, and the percentage of exercises from each source that were labeled as reasoning-and-proving by the coder. It is notable that of the exercises taken or adapted from ancillary resources or created by Karen, 41 out of 49 (83.7%) were reasoning-and-proving exercises, indicating that when Karen did stray from her textbook, she choose exercises that would engage her students in reasoning-and-proving.

Table 4.4 Sources of Exercises Selected by Karen

Source	**All Exercises Frequency**	**Reasoning-and-Proving Exercises Frequency**
Taken Directly from Published Textbook/Curriculum	259	188 (72.6%)
Modified from Textbook/Curriculum	27*	27 (100%)
Used in the CORP Course	0	0 (0 %)
Taken or Adapted from Ancillary Resources	38	32 (84.2%)
Created by Teacher	11	9 (81.8%)

*An additional 6 exercises were labeled as modified by Karen on her Task Log Sheet for Day 7, but the modifications were just reminders to students to draw pictures, decide the truth or falsity of a statement, and back up their decisions with an argument or counterexample. These were the same directions that were with the exercises in the textbook. The "modification" label was removed from these 6 items by the primary researcher.

4.1.2 Implementation of reasoning-and-proving exercises *(RQ2: To what extent were participants able to maintain the level of cognitive demand of reasoning-and-proving tasks during implementation?)*

All of the exercises chosen by Karen for her student work samples had the potential to engage her students in proof (see Figures 4.4, 4.5, and 4.6). In the IQA Rubric for Potential of the Task, a task that scores 4 "must explicitly prompt for evidence of students' reasoning and understanding. For example, the task MAY require students to…create a proof or find a counterexample" (see Appendix K). An implementation score of 4 indicates that "there is explicit evidence of students' reasoning and understanding. For example, students may have…created a proof or found a counterexample" (see Appendix K). Karen's first classroom artifact task asked students to create their first proof; while not every sample of student work that Karen labeled "exceeded" or "met" expectations was rated as proof by the coders, the students' samples showed that Karen maintained a high level of cognitive demand for the task. According to Boston and Smith (2009), score levels 3 and 4 on the IQA Rubrics for Potential of the Task and Implementation of the Task "represent high-level cognitive demands in which the connections to meaning and understanding are implicitly (score level 3) or explicitly (score level 4) required by the task” (p.153).

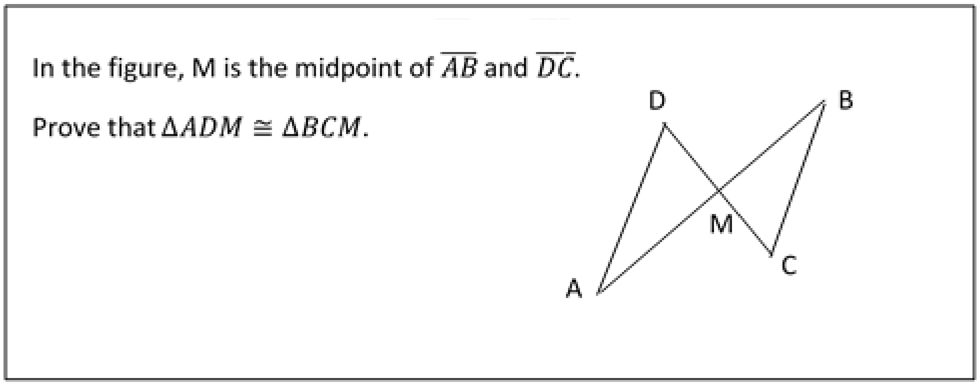


Figure 4.4 First classroom artifact packet task assigned by Karen (Day 1).

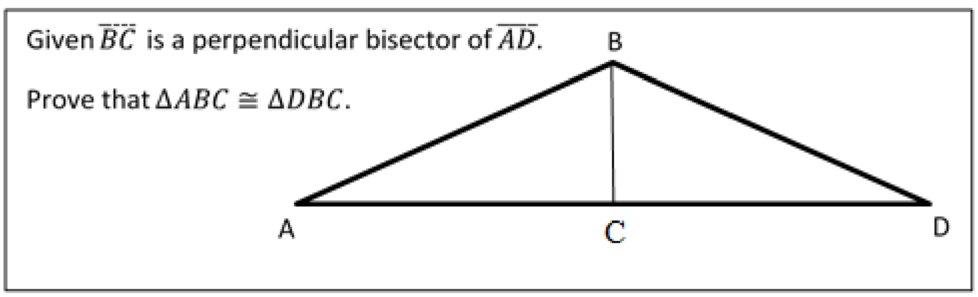


Figure 4.5 Second classroom artifact packet task assigned by Karen (Day 7).

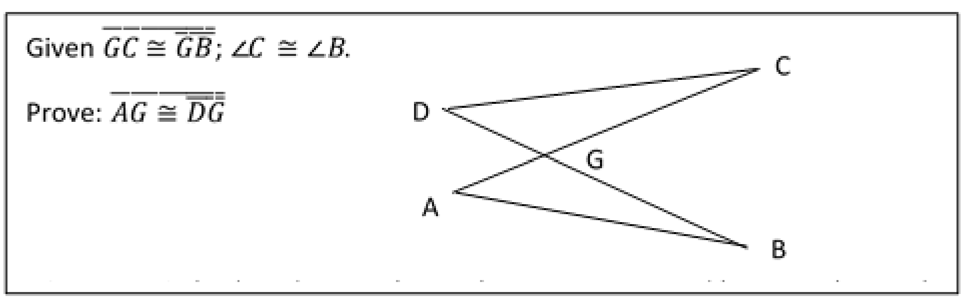


Figure 4.6 Third classroom artifact packet task assigned by Karen (Day 11).

Figure 4.7 shows an example of a piece of student work for the third classroom artifact packet task (Figure 4.6). The proof is complete and is essentially in a two-column format; the task required a proof and this student ("Karen 3-1") produced a proof. Karen's third classroom artifact packet contained a collection of three exercises; Karen labeled this student's work as "exceeding expectations." Karen maintained the level of cognitive demand for all of her classroom artifact packet tasks. Table 4.5 summarizes the IQA rubric scores for the potential of the task and the implementation of the task for Karen's classroom artifact packets. All of Karen's scores were 3 or 4, indicating that she successfully chose tasks with the potential to engage her students in reasoning-and-proving and she maintained the level of cognitive demand through the implementation of those tasks.

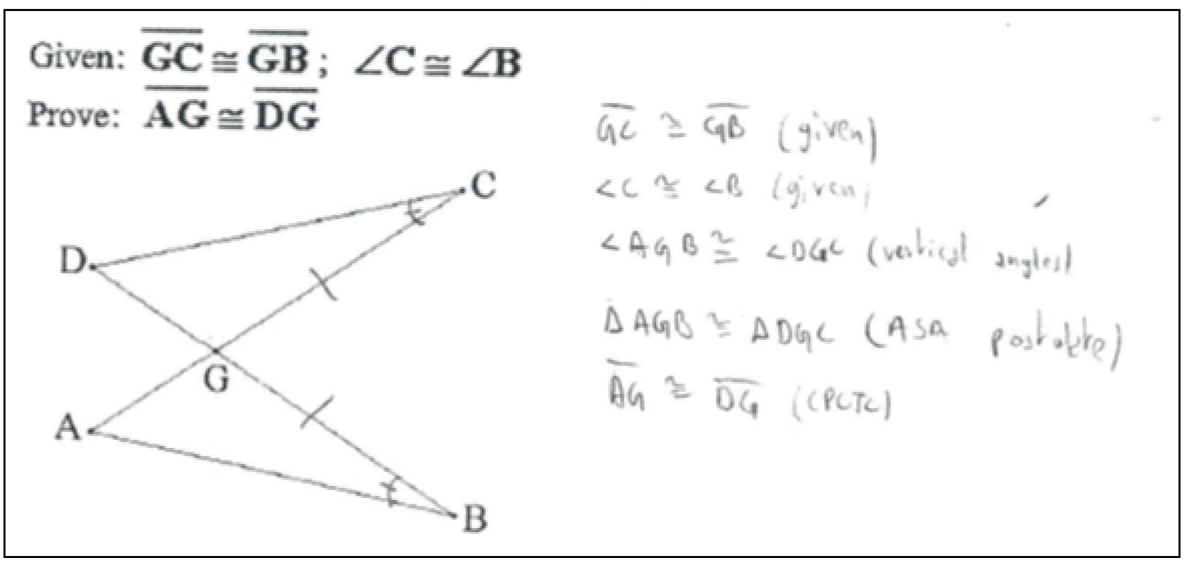


Figure 4.7 Classroom artifact sample Karen 3-1 showing a valid proof.

Table 4.5 Potential and Implementation of Karen's Tasks for Cognitive Demand

Classroom Artifact Packet	Potential of Task	Implementation of Task	Maintained Cognitive Demand?
First	4	3	Maintained
Second	4	4	Maintained
Third	4	4	Maintained

4.1.3 Evaluation students' reasoning-and-proving products ***(RQ3: To what extent were participants able to accurately evaluate their students' reasoning-and-proving products?)***

The core elements of proof used in the reasoning-and-proving course in which Karen, Uma, Sidney, and Jonathan participated were:

- The argument must show that the conjecture or claim is (or is not) true for *all* cases.
- The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.
- The conclusion that is reached from the set of statements must follow logically from the arguments made.

When asked how she would define proof to a colleague, Karen responded, "Make a claim and list statements with accompanying justification in a logical order to prove your goal statement" (personal communication, May 27, 2013). When asked how she would define proof to a student, Karen responded slightly differently:

> [Proving is when you construct] a valid argument where you are listing what you are given, where you use what you know (e.g., definitions and theorems) to reach a goal.

> The goal is a statement that you are trying to prove. For the proof, each statement you list has to be backed up (justification). A reader has to be able to follow your argument because the reader is not inside your head; you have to put down enough detail so that the reader can understand your process and argument. (personal communication, May 27, 2013)

Comparing these definitions to the core elements of proof (see Table 4.6), Karen captured using true and validated statements to justify claims and the logical flow of the argument, but she was vague about the "goal statement." It is possible that she understands a "goal statement" must apply to all cases in a particular domain, but she did not voice that either initially or under probing.

Table 4.6 Comparison of the Core Elements of Proof with Karen's Definition of Proof

Core element of proof	**Karen**
The argument must show that the conjecture or claim is (or is not) true for *all* cases.	Partial
The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.	Present
The conclusion that is reached from the set of statements must follow logically from the argument made.	Present
Additional Criteria	None

4.1.3.1 Criteria for judgment ***(RQ3a: To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?)***

Karen's expectations for her students included the idea of "formal proof," although she did not define proof until her post-interview. Based on that information and the classroom artifact packet tasks themselves, it is clear that Karen expected her students to show that a particular conjecture was true for all figures with the given characteristics, the statements and definitions used must have been defined by the classroom community (students keep a list of these in tan composition books in Karen's classroom), and the conclusion reached by the students must follow logically from the argument made (core elements of proof). In addition, Karen delineated level of quality. For example, the rubric for the first sample student work was described by Karen as follows (Karen's first classroom artifact packet Cover Sheet):

- "Exceeded expectations" for students who wrote a formal proof with no errors
- "Met expectations" for students who stated and justified some claims but not all, or made some errors in terminology but overall were able to justify their claims; note that this is the first proof the students had attempted
- "Failed expectations" for students who wrote a postulate or made assumptions without justification

Karen's rubric for the second student work exercise was virtually identical to the first rubric (there were some minor changes in wording in "met" and "failed" expectations).

The third rubric covered three exercises on a quiz, was holistic for the entire quiz, and was based on completeness and correctness (see Table 4.7). A score of 4 on the IQA rubric for Clarity and Detail of Expectations means that "the expectations for the quality of students' work are very clear and elaborated. Each dimension or criterion for the quality of students' work is clearly articulated. Additionally, varying degrees of success are clearly differentiated" and proof evaluations must be based on the core elements (see Appendix K). The raters agreed that Karen scored a 4 (highest level) on all three tasks for her clarity and detail of expectations (see Table 4.8).

Table 4.7 Karen's Rubrics for Her Classroom Artifact Packets

<table>
<tr><td>Rubric Level</td><td>First Rubric</td><td>Second Rubric</td><td>Third Rubric</td></tr>
<tr><td>Exceeded expectations</td><td>Formal proof with no errors</td><td>Formal proof with no errors</td><td rowspan="3">Mastery:
10 points: no errors

8-9 points: one or two small errors while writing out the proof, but not an error in logic

Not Mastered:
6-7 points: Wrote a proof for one of the problems but not both (including justification) OR wrote statements and reasons for parts of both proofs but did not complete either accurately

5 points or below: Incomplete, little or no correct statements with justification</td></tr>
<tr><td>Met expectations</td><td>Stated and justified some claims but not all, or made some errors in terminology but overall were able to justify their claims</td><td>Formal proof but with one or two minor errors</td></tr>
<tr><td>Failed expectations</td><td>Wrote a postulate or made assumptions without justification</td><td>Assumed that segments and angles were congruent without justification or did not provide justification for statements</td></tr>
</table>

Table 4.8 Quality of Karen's Rubrics Used to Judge Student Work

Classroom Artifact Packet Rubric	Clarity and Detail of Expectations	Comment
First	4	*Used core elements of proof and levels of expectations*
Second	4	*Used core elements of proof and levels of expectations*
Third	4	*Used core elements of proof and levels of expectations*
Average	4	

4.1.3.2 Application of core elements of proof (*RQ3b: To what extent did participants apply the core elements of proof in evaluating their students' reasoning-and-proving products?)*

Karen agreed with the coders in identifying which of her students' reasoning-and-proving products qualified as proof with one exception: student sample Karen 1-1 (see Figure 4.8). At first glance, this students' work looks complete and logical; it is easy to miss that the student mislabeled one of the angles. When this was discussed with the second coder, the second coder felt that the error was minor and since the student had correctly marked the drawing, the work still qualified as proof. The primary researcher asked Karen to render an opinion about this error, and while she missed it the first time she scored the work, Karen agreed with the primary researcher that the work—while still exceeding expectations for the student's first proof attempt—did not actually qualify as a proof due to the error. Karen explained that she did not expect students to list the given statements in this first proof attempt (personal communication, May 27, 2013). Student work sample Karen 1-2 did qualify as a proof because it contained the core elements of proof, showed no errors, and included the given statements (see Figure 4.9).

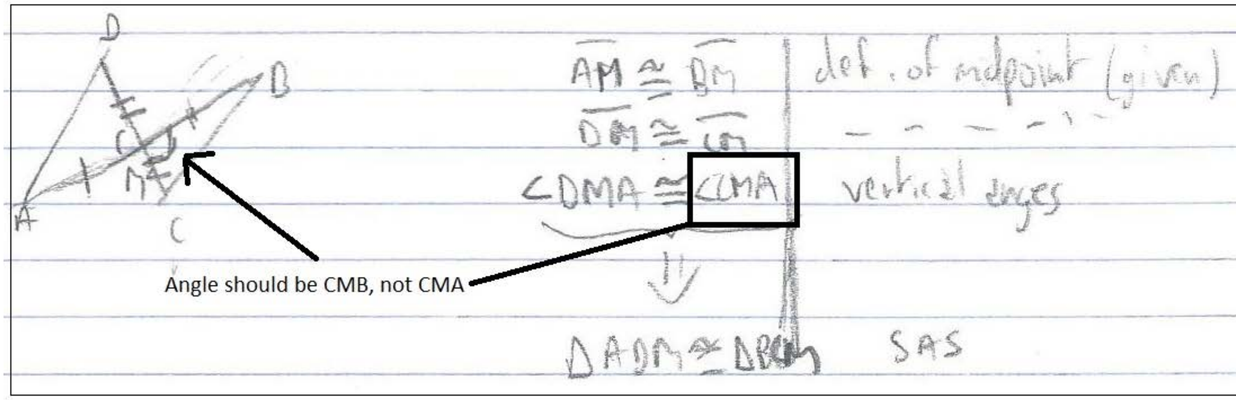


Figure 4.8 Classroom artifact packet sample Karen 1-1 showing a non-proof argument due to the mislabeling of an angle.

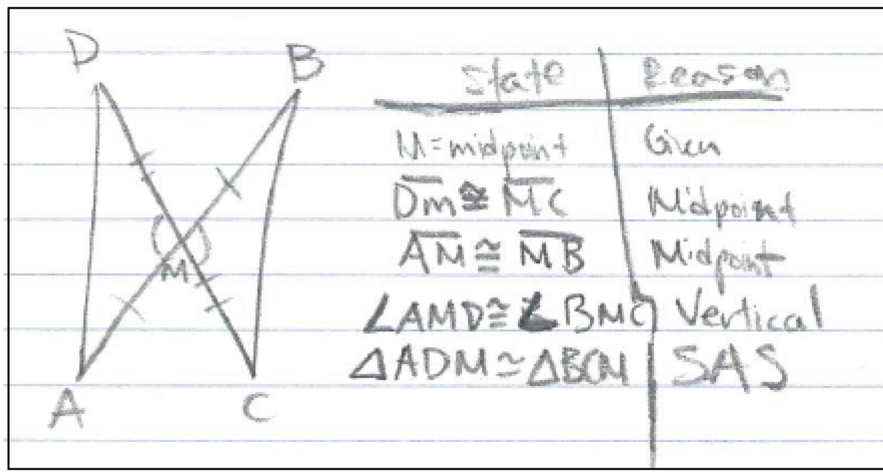


Figure 4.9 Classroom artifact packet sample Karen 1-2 showing a valid proof rated "exceeded expectations".

The scores of all of the student work samples in the classroom artifact packets are displayed in Table 4.9. As mentioned before, Karen did not completely base her expectation score on the core elements of proof; she judged expectations on the developmental progress of her students. The information contained in the column for whether or not Karen judged each sample of student work to be a proof was obtained during her post-interview on May 27, 2013, which was conducted in person so Karen could review the samples of student work she had collected.

Table 4.9 Karen's Application of the Core Elements of Proof in Evaluating Student Work

Classroom Artifact Packet Sample	Participant Score	Participant Evaluation: Proof?	Researcher Evaluation: Proof?	Comments
First task: Prove that two triangles are congruent (first solo proof exercise attempted by students; see Figure 4.4)				
Karen 1-1	Exceeded expectations	Nonproof	Nonproof	*Given statement missing; angle mislabeled in vertical angles (should be angle DMA is congruent to angle CMB)*
Karen 1-2	Exceeded expectations	Proof	Proof	
Karen 1-3	Met expectations	Nonproof	Nonproof	*Did not state which segments were congruent in two-column proof; congruent sides were labeled in drawing*
Karen 1-4	Met expectations	Nonproof	Nonproof	*Stopped after listing given and two pairs of congruent sides; marked congruent vertical angles in drawing*
Karen 1-5	Failed expectations	Nonproof	Nonproof	*Series of incorrect claims; did state SAS but without coherent justification except to label two pairs of congruent sides in drawing*
Karen 1-6	Failed expectations	Nonproof	Nonproof	*Drawing correctly labeled and SAS stated but no series of statements and justifications*
Second task: Prove that two triangles are congruent (see Figure 4.5).				
Karen 2-1	Exceeded Expectations	Proof	Proof	*Statements and justifications used are those accepted by the community but needs to justify the first statement as "given"*

Karen 2-2	Exceeded Expectations	Proof	Proof	
Karen 2-3	Met Expectations	Nonproof	Nonproof	*Did not state the given statement*
Karen 2-4	Met Expectations	Nonproof	Nonproof	*First statement is unclear—might say "$\overline{BC}$ is shared" implying $\overline{BC}$ is congruent to itself by the reflexive property; include the given statement; replace "they" in the fourth statement with specific triangles*
Karen 2-5	Failed Expectations	Nonproof	Nonproof	*No justifications; argument based on SSS triangle congruence which is not possible here*
Karen 2-6	Failed Expectations	Nonproof	Nonproof	*No mention that AC = CD and an incorrect justification for the congruent angles.*

Third task: Prove that two corresponding parts of two (congruent) triangles are congruent (see Figure 4.6).

Karen 3-1	Exceeded Expectations	Proof	Proof	
Karen 3-2	Exceeded Expectations	Proof	Proof	
Karen 3-3	Met Expectations	Proof	Proof	
Karen 3-4	Met Expectations	Proof	Proof	*Minor error by not labeling the vertical angles with three points (only used the vertex)*
Karen 3-5	Failed Expectations	Nonproof	Nonproof	*No justifications or organization; SSS was incorrectly used as the type of triangle congruence*

Karen 3-6	Failed Expectations	Nonproof	Nonproof	*Congruent parts of the triangles were labeled on the diagram but no justifications or type of triangle congruence was provided*

When initially examining the second set of student work samples, the primary researcher thought that students had not properly dissected the implications of the given statement, "$\overline{BC}$ is a perpendicular bisector of $\overline{AD}$." She expected the proofs to include a series of statements and justifications (or some of them) such as shown in Figure 4.10. When asked why her students did not include these lines in their proofs, Karen showed the primary researcher the definition of "perpendicular bisector" as recorded by students in their tan composition notebooks: "A line perpendicular to a segment's midpoint" (from a student's notes). The definition was accompanied by a picture that showed a triangle split by a perpendicular bisector with two right angles and two labeled, congruent segments on either side of the midpoint. Not every perpendicular bisector splits a triangle in half (this only works for isosceles and equilateral triangles), but the labeled picture in the students' notes matched the conditions in the triangle in the second classroom artifact packet task, so students could take as shared by their community that a perpendicular bisector with this configuration (segment passes through the opposite vertex) creates two congruent right angles and two congruent segments on either side of the midpoint. Thus, the steps expected by the primary researcher were not necessary for proof in Karen's class. Figure 4.11 shows a valid proof and Figure 4.12 shows an insufficient argument for the second student work sample proof.

Statement	Justification
$\overline{BC}$ is a perpendicular bisector of $\overline{AD}$.	Given
$\angle ACB$ and $\angle DCB$ are right angles	Definition of perpendicular segments
$m\angle ACB = 90°$ and $m\angle DCB = 90°$	Definition of right angles
$m\angle ACB = m\angle DCB$	Transitive property or substitution
$\angle ACB \cong \angle DCB$	Definition of Congruent Angles
C is a midpoint of $\overline{AD}$	Definition of bisector
$\overline{AC} \cong \overline{CD}$	Definition of midpoint

Figure 4.10 Expected lines in the beginning of the second classroom artifact packet proof

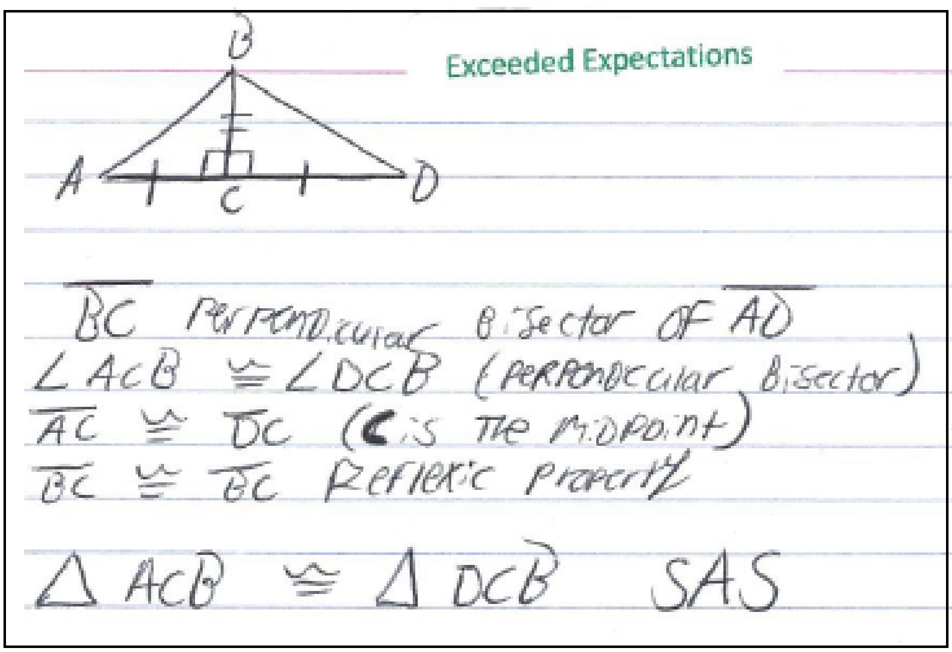


Figure 4.11 Second classroom artifact packet sample Karen 2-1 showing a proof.

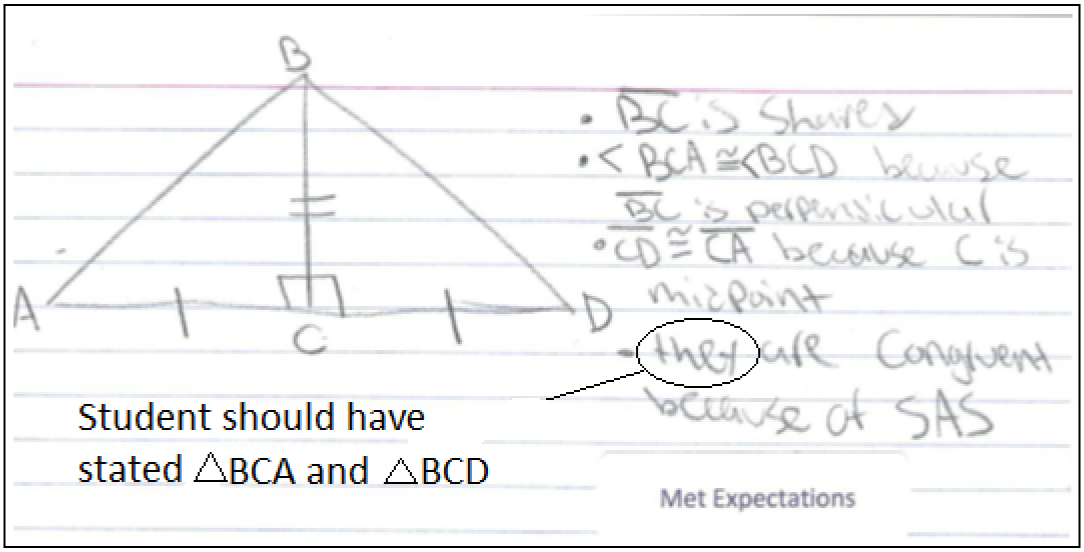


Figure 4.12 Second classroom artifact packet sample Karen 2-4 showing an insufficient argument.

4.1.3.3 Communication of expectations *(RQ3c: In what ways did participants communicate expectations regarding what is required to produce a proof to students?)*

One way to assess how a teacher communicates expectations to students is to ask the students. On Day 11 of Karen's data collection period, she gave the students a mastery assessment so that she could adjust her lesson plans. Her first question was, "why is proof so important in mathematics?" The responses from six students were as follows (Karen's comments are in italics):

- "To obtain accurate results in problem solving by explaining the logic that went into a solution" (Karen 3-1)
- "So you can prove that something is true. You can prove that your answer works. If you didn't prove then you could say anything you wanted." *Great point! Justification is key!*

(Karen 3-2)

- "Proof is so important because it extends the knowledge of a subject, based on a proven previous subject. Also that allow for further explanation into a theorem." *Nice! Why is justification such an important part of proof writing?* (Karen 3-3)
- "If someone just gives an answer with no proof then it is basically useless for other problems. If someone prooves [sic] something once then they know it is true for the next and future problems." *Great point! We often use things that we prove as justification in other proofs!* (Karen 3-4)
- "Proof is important in mathematics because without it, yourself and nobody else will actually know if your answer is right." *I agree. Why is justification so important?* (Karen 3-5)
- "To get the right answer and to find who is right or wrong." *Absolutely! Why is providing justification so important?* (Karen 3-6)

The students' comments indicate that Karen conveyed the idea that proof is a logical argument used to develop a verified body of knowledge which can be applied or used to further mathematical understanding. Karen's notes on the classroom artifact packet cover sheets also indicate that she discussed her expectations for student work and her lesson plans indicated that students saw many examples of proofs. By the third classroom artifact packet task, students were given written directions to "prove the following statements, using any of the styles and methods you have learned" and were reminded orally to justify and organize their arguments in a way that their audience could understand clearly. The IQA rubric for Communication of Expectations states that teachers who discuss their expectations or criteria for student work with students in advance of their completing the assignment and model high-quality work are rated 4 on the rubric (see Appendix K). When shown the IQA rubric for Communication of Expectations for the purpose of member checking, Karen agreed with the primary researcher and the second coder that she should be rated a 4 (highest level) for each task (see Table 4.10). In addition, Karen reinforced her expectations by the feedback she provided her students on their work (see Figure 4.13).

Table 4.10 Karen's Communication of Expectations Rubric Scores

Classroom Artifact Packet	Communication Score
First	4
Second	4
Third	4
Average Score	4

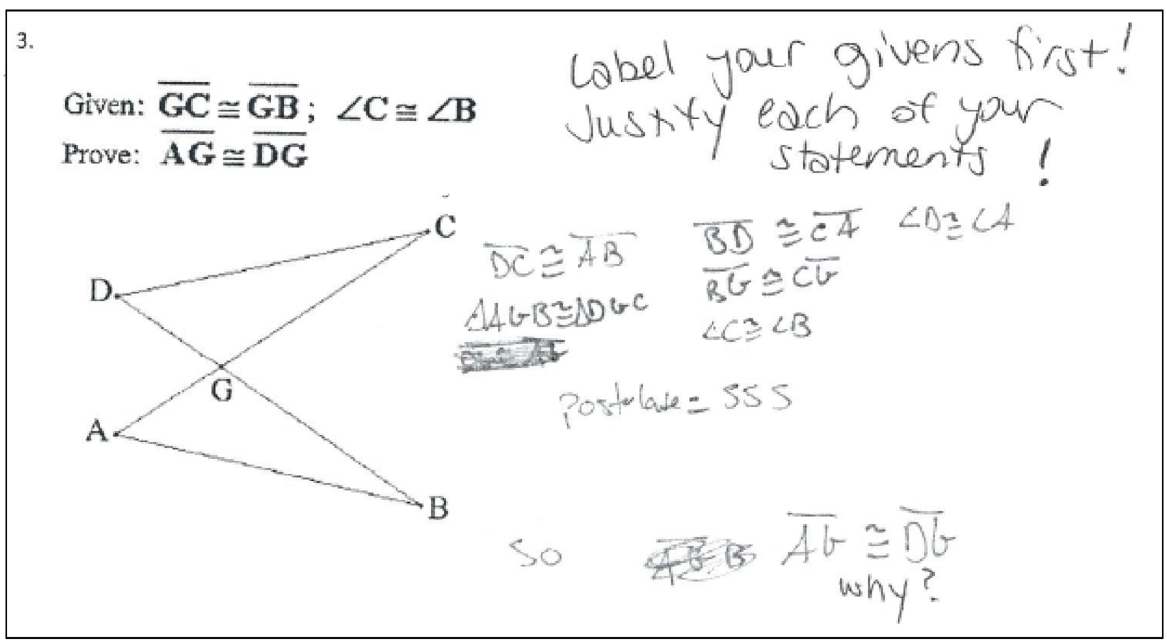


Figure 4.13 Classroom artifact packet sample Karen 3-5 showing an insufficient argument and Karen's comments.

In conclusion, Karen's district uses a reform curriculum that supports reasoning-and-proving to a high degree in the geometry course. For the pages of CME's *Geometry*, Unit 2 that Karen covered during her 15-day data collection period, 75.5% of the exercises had the potential to engage students in reasoning-and-proving. Of the 335 exercises Karen selected from all sources, Karen offered her students 256 opportunities to engage in reasoning-and-proving (76.4%). About two-thirds of the reasoning-and-proving exercises Karen selected were of the types Investigate a Claim or Develop an Argument. While Karen did not modify any exercises to include more reasoning-and-proving, she did not lower the potential of any of the exercises. Although she mainly used exercises from her CME *Geometry* curriculum, when Karen did adopt from ancillary sources or modify curriculum tasks, she choose reasoning-and-proving exercises. Finally, Karen's student work samples indicate that she can recognize and implement a cognitively-demanding task, construct clear and detailed expectations based on the core elements of proof, evaluate students' work based on the core elements, and communicate those expectations to her students.

4.2 UMA

Uma also teaches geometry in an urban district which has a racially mixed population of students with low socio-economic status. Unlike Karen, though, Uma teaches in Virginia which is a state that did not adopt the Common Core State Standards and places very little emphasis on reasoning-and-proving. For example, the district pacing guide given to Uma for her geometry class does not include triangle congruence proofs, a mainstay of geometry proofs (personal communication, May 29, 2013). To make sure teachers follow the district pacing guide, students are given a district-created benchmark assessment every nine weeks and the Virginia Standards of Learning (SOL) test at the end of the year. Uma reports that about half of her teaching evaluation is based on her students' performance on the SOLs.

Uma's school is one of four high schools in her district and has about 2,000 students in grades 9-12. Uma reported that most of the other mathematics teachers in her school are experienced and teach in a traditional manner that does not resonate with her university training. Since the district insists that all teachers with the same subject participate in Professional Learning Communities (PLC), follow the same pacing guide, and administer common exams, Uma has difficulty assessing her students in the way that she would like (e.g., open-ended questions). Uma teaches two geometry classes and two Algebra 2 inclusion classes. She selected a geometry class for this study because she taught geometry last year and had a sense of where she could include reasoning-and-proving tasks and still follow the district pacing guide.

4.2.1 Selecting exercises *(RQ1: To what extent did participants select reasoning-and-proving learning opportunities in the form of exercises?)*

4.2.1.1 Available in textbook *(RQ1a: To what extent does the textbook include exercises that have the potential to engage students in reasoning-and-proving?)*

The textbook used by Uma's class was Glencoe's *Geometry*, Virginia edition (Boyd, Cummins, Mallow, Carter, & Flores, 2005). In the fifteen days of data collection, Uma's class studied areas of circles, quadrilaterals and polygons and surface area of solids (chapters 11 and 12 in the textbook). The textbook exercises consisted of guided practice, practice and apply, mixed review, getting ready for the next lesson, practice quizzes, chapter study guides and reviews, and practice chapter tests. There were 783 exercises spanning area and surface area; 41 one of them were coded as reasoning-and-proving exercises which represents 5.2% of the available exercises.

Uma's textbook exercises favored making generalizations rather than developing arguments. Over three-quarters of the exercises asked students to make or investigate conjectures; less than a quarter of the exercises asked students to develop an argument or find a counterexample (see Table 4.11). As a reminder, some exercises were coded as more than one type so the sum of the exercises listed in Table 4.11 does not sum to 41.

Table 4.11 Available Types of Reasoning-and-proving Exercises in Uma's Textbook

Type of RP Exercise	Available in Textbook*
Make a Conjecture	15 (1.9%)
Investigate a Conjecture	18 (2.3%)
Evaluate an Argument	4 (0.5%)
Correct a Mistake	0 (0%)
Develop an Argument	5 (0.6%)
Counterexample	2 (0.3%)
Principles of Proof	0 (0%)
Non-Reasoning-and-Proving	743 (94.8%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

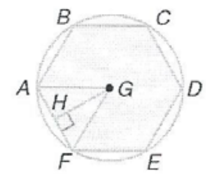
4.2.1.2 Selected by Teacher *(RQ1b: To what extent did the participant select exercises for instruction that had the potential to engage students in reasoning-and-proving?)*

Uma's lessons included warm-up exercises, some exploratory and practice exercises, chapter reviews, and exams. In the 15-day data collection period, Uma assigned 291 exercises to her students, 10 of which were reasoning-and-proving exercises (3.0%), which is lower than the percentage contained in the textbook (5.2 %). In her interviews, Uma cited time constraints as the reason she did not engage her students more in reasoning-and-proving. In her district, all geometry teachers must give the same chapter exams which model the Virginia SOL exam and are written by experienced teachers not focused on reasoning-and-proving. In order to prepare her students for their exams, Uma feels that she must devote much of her teaching to conveying procedures instead of offering opportunities to engage students in reasoning-and-proving.

Of the 10 exercises assigned by Uma that the coders judged to be reasoning-and-proving, Uma labeled 9 of them to be reasoning-and-proving and one not reasoning-and-proving (a claim investigation from the chapter review game created on Day 13). Of the 281 exercises selected by Uma that were not rated as reasoning-and-proving by the coders, Uma labeled 32 as reasoning-and-proving and 249 as not reasoning-and-proving. This analysis may under- or over-represent Uma's ability to recognize reasoning-and-proving tasks because Uma collectively gave large groups of exercises one label on her Task Log Sheets. For example, Uma labeled the entire Day 13 chapter review game as "not reasoning-and-proving," and indeed 23 of the 24 exercises were not, but one asked students to investigate a conjecture. Similarly, Uma considered the exercises shown in Figure 4.14, Figure 4.15, and 12 practice exercises from Day 4 as a set and wrote that "some questions" were reasoning-and-proving. In contrast, the coders thought that the scaffolding and the lack of other regular polygons prevented the first four exercises in Figure

4.14 from rising to the level of reasoning-and-proving. The fifth exercise, however ("What is the relationship between the perimeter of the regular hexagon and its area?") was rated as reasoning-and-proving because students were asked to make a conjecture.

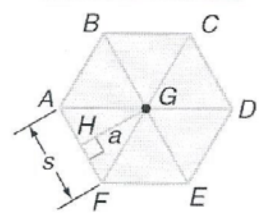
Describe the figure to the right.



Apothem: ______________________________

1. Polygon $ABCDEF$ is a regular ______________. What do you know about all regular polygons?

2. What is the relationship between the area of ΔAGF and the area of the hexagon?

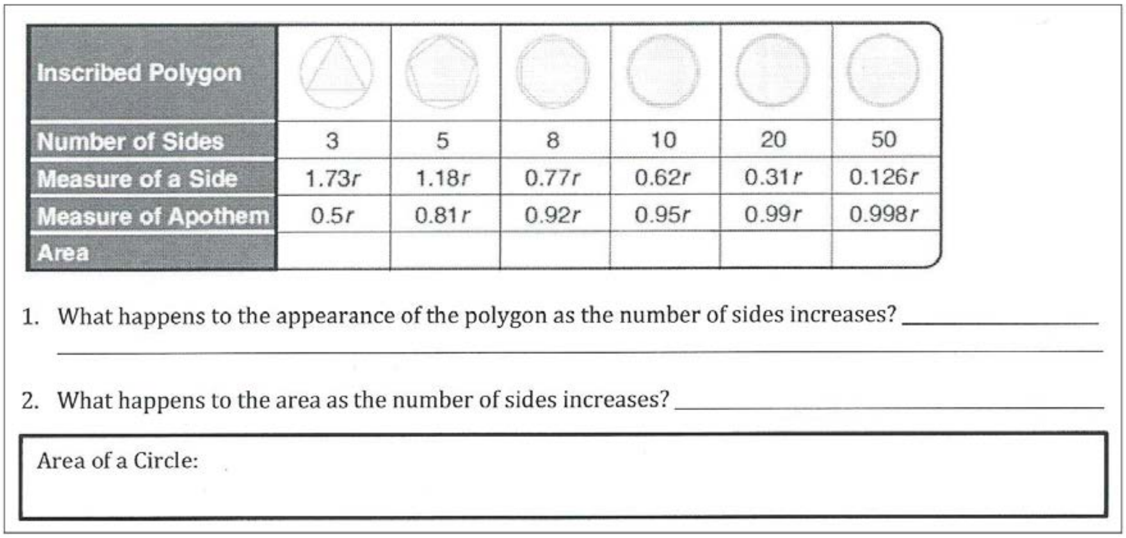


3. In the figure to the right, ***GH*** is an ______________ which is ***a*** units long.
All sides of the hexagon are _______ units long.
Find the area of ΔAGF. Use this to find the area of the hexagon.

4. What is the perimeter of the regular hexagon? ____________
5. What is the relationship between the perimeter of the regular hexagon and its area?

Area of a Regular Polygon:

Figure 4.14 Developing the formula for the area of a regular polygon (Uma's Lessons, Day 4).



Inscribed Polygon						
Number of Sides	3	5	8	10	20	50
Measure of a Side	1.73r	1.18r	0.77r	0.62r	0.31r	0.126r
Measure of Apothem	0.5r	0.81r	0.92r	0.95r	0.99r	0.998r
Area						

1. What happens to the appearance of the polygon as the number of sides increases? ________________
2. What happens to the area as the number of sides increases? ________________

Area of a Circle:

Figure 4.15 Developing the formula for the area of a circle (Uma's Lessons Day 4; Boyd et al., 2005, p. 611).

Uma's Day 4 lesson also focused on the area of a circle. This task used in this lesson was taken from her textbook, and it shows a pattern of polygons with increasing number of sides. From the information provided to students in the table, students can make a conjecture about the formula for the area of a circle. Because of the pattern, the two exercises in this task were rated as reasoning-and-proving (see Figure 4.15).

Describing Uma's reasoning-and-proving engagement opportunities in terms of the number of exercises she offered may under-represent what her students experienced. A more accurate representation may be to look at the time her students spent on reasoning-and-proving exercises. On Day 1, Uma's students spent part of 40 minutes developing the formulas for the areas of parallelograms, trapezoids, and rhombi. On Days 3 and 4, students spent a total of 40 minutes proving or disproving the statement "If two polygons have the same area, then they are congruent" (this was also Uma's first classroom artifact packet task). On Day 4 students also spent part of 50 minutes developing formulas for areas of regular polygons and circles (see

Figures 4.14 and 4.15). On Day 8 students spent the entire 90 minute class on a fencing task (third classroom artifact packet task), and on Day 11 students spent 20 minutes developing a formula for the surface area of a regular pyramid. If the time spent on reasoning-and-proving on Day 1 (part of 40 minutes) and Day 4 (part of 50 minutes) is estimated to be 20 minutes and 25 minutes, respectively, then students spent an estimated 195 minutes out of 1215 minutes of class time over 15 days potentially engaging in reasoning-and-proving. This measure-16 % of class time—indicates a much higher focus on reasoning-and-proving than the number of exercises suggests.

Similar to the textbook, more of Uma's reasoning-and-proving exercises focus on making generalizations than on developing arguments; 10 exercises involved making or investigating a conjecture, while 6 exercises asked students to develop an argument or find a counterexample (6 of the exercises were coded for more than one action). Table 4.12 provides the details.

Table 4.12 Uma's Selected Types of Reasoning-and-proving Exercises

Type of RP Exercise	Selected by Teacher*
Make a Conjecture	8 (2.8%)
Investigate a Conjecture	2 (0.7%)
Evaluate an Argument	0 (0%)
Correct a Mistake	0 (0%)
Develop an Argument	5 (1.7%)
Counterexample	1 (0.3%)
Principles of Proof	0 (0%)
Non-Reasoning-and-Proving	274 (94.5%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

In comparing the available exercises in the textbook with the exercises that Uma offered her students (see Figure 4.16), it is clear that Uma offered more Make a Conjecture exercises (2.8% compared to 1.9%) and Develop an Argument exercises (1.7% compared to 0.6%) than were available in the textbook. Almost all of the Develop and Argument exercises were created by Uma (one for the Student Work packets and three in Day 1 Lessons about developing formulas for the areas of parallelograms, trapezoids, and rhombi).

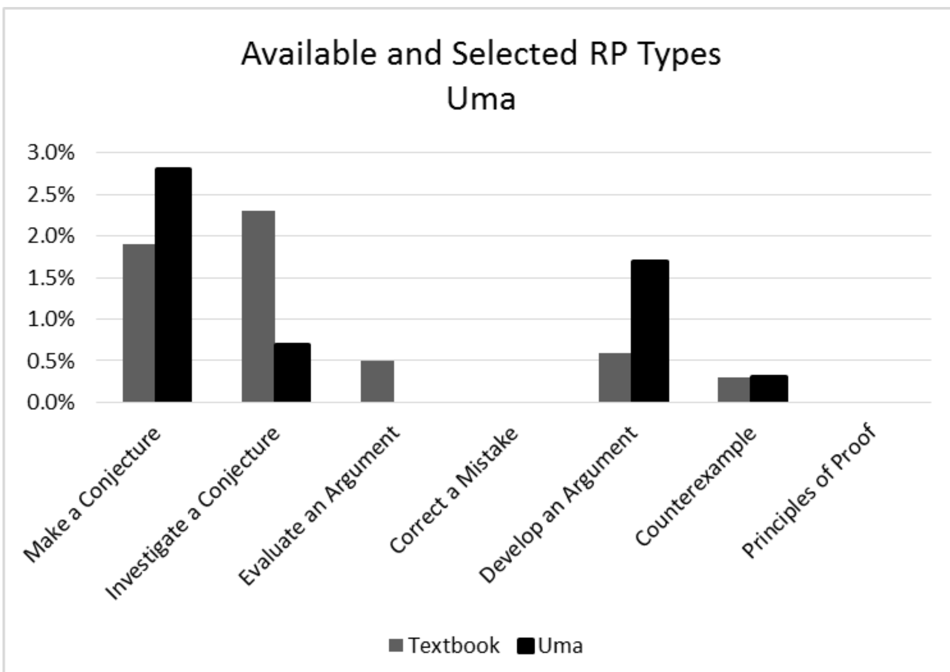


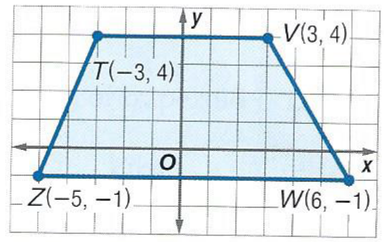
Figure 4.16 Comparison of available and selected types of reasoning-and-proving exercises for Uma

4.2.1.3 Exercise modifications *(RQ1c: To what extent did the participant modify tasks to affect the tasks' potential to engage students in reasoning-and-proving?)*

According to Uma, virtually all of her modifications involved simplifying the wording of a problem or adding extra questions, such as asking students to identify a shape, write down a formula, or explain their thinking. Figure 4.17 shows an example from Uma's textbook that she modified to include extra practice for her students with vocabulary and formulas (see Figure 4.18 for the modified exercises). In total, Uma labeled 131 exercises as "modified," but this label was applied to collections of exercises, not individual exercises. Regardless, only 3 modified exercises had the potential to engage students in reasoning-and-proving, and this section will focus on these three exercises. In the first modified exercise, Uma took a textbook presentation of the development of the formula for the area of a regular polygon and modified it into the collection of exercises shown in Figure 4.14. The fifth question is a make-a-conjecture exercise, and is coded as a non-reasoning-and-proving exercise modified to include reasoning-and-proving. Uma made a similar modification in her lesson from Day 11 for the surface area of a regular pyramid (see Figure 4.19 for the textbook presentation and Uma's modification in Figure 4.20). The third modified reasoning-and-proving exercise is the claim investigation from the chapter review on Day 13. This exercise, which asks students to determine if the inverse of the statement "if a polygon has four sides then it is a quadrilateral" is true or false is similar to exercises found in Uma's textbook on p. 79. Those exercises were already reasoning-and-proving exercises, so Uma's wording modification has a neutral effect. This data is summarized in Table 4.13.

Example 2 Area of a Trapezoid on the Coordinate Plane

COORDINATE GEOMETRY **Find the area of trapezoid *TVWZ* with vertices $T(-3, 4)$, $V(3, 4)$, $W(6, -1)$, and $Z(-5, -1)$.**



Bases: Since $\overline{TV}$ and $\overline{ZW}$ are horizontal, find their length by subtracting the x-coordinates of their endpoints.

$TV = |-3 - 3| = |-6|$ or 6 $\qquad ZW = |-5 - 6| = |-11|$ or 11

Height: Because the bases are horizontal segments, the distance between them can be measured on a vertical line. That is, subtract the y-coordinates.

$h = |4 - (-1)|$ or 5

Area:

$A = \frac{1}{2}h(b_1 + b_2)$ Area of a trapezoid

$= \frac{1}{2}(5)(6 + 11)$ $h = 5, b_1 = 6, b_2 = 11$

$= 42.5$ Simplify.

The area of trapezoid *TVWZ* is 42.5 square units.

Figure 4.17 Original exercise: Determining the area of a trapezoid (Boyd et al., 2005, p. 603).

Plot points A(1,-2), B(5, -2), C(4,4), D(1,4). What type of shape is this?
Explain how you know ______________________________________

Write the formula and find the area of quadrilateral ABCD.

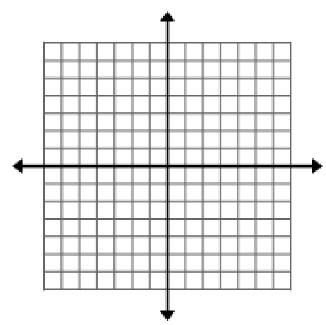
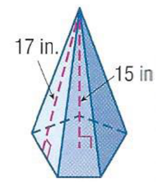


Figure 4.18 Modified exercise: Determining the area of a trapezoid (Uma's Lesson Day 2).

Example 3 Surface Area of Pentagonal Pyramid

Find the surface area of the regular pyramid.



The altitude, slant height, and apothem form a right triangle. Use the Pythagorean Theorem to find the apothem. Let a represent the length of the apothem.

$c^2 = a^2 + b^2$ Pythagorean Theorem

$(17)^2 = a^2 + 15^2$ $b = 15, c = 17$

$8 = a$ Simplify.

Now find the length of the sides of the base. The central angle of the pentagon measures $\frac{360°}{5}$ or 72°. Let x represent the measure of the angle formed by a radius and the apothem. Then, $x = \frac{72}{2}$ or 36.

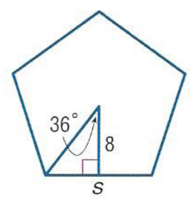
Use trigonometry to find the length of the sides.

$\tan 36° = \frac{\frac{1}{2}s}{8}$ $\tan x° = \frac{\text{opposite}}{\text{adjacent}}$

$8(\tan 36°) = \frac{1}{2}s$ Multiply each side by 8.

$16(\tan 36°) = s$ Multiply each side by 2.

$11.6 \approx s$ Use a calculator.



Next, find the perimeter and area of the base.

$P = 5s$

$\approx 5(11.6)$ or 58

$B = \frac{1}{2}Pa$

$\approx \frac{1}{2}(58)(8)$ or 232

Finally, find the surface area.

$T = \frac{1}{2}P\ell + B$ Surface area of a regular pyramid

$\approx \frac{1}{2}(58)(17) + 232$ $P \approx 58, \ell = 17, B \approx 232$

≈ 726.5 Simplify.

The surface area is approximately 726.5 square inches.

Figure 4.19 Original exercise: Determining formulas for the lateral and surface areas of regular pyramids (Boyd et al., 2005, p. 662)

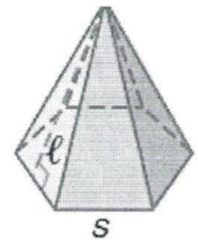
Base ____

Vertex ____

Lateral Faces ____

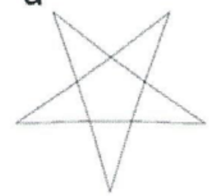
Slant Height ____

Regular Pyramid ____



The figure is a regular ____________________. The perimeter of the base is a _________
Each lateral face is a ______________________ with base ________ and height ________
Find the lateral area. How does this relate to the perimeter?

The base of the pyramid is a _______________________
The area of the base is____
The surface area of a pyramid is the sum of the _______________ and the area of the ______
Find the surface area of the pyramid.



Lateral Area of a Regular Pyramid	Surface Area of a Regular Pyramid

Figure 4.20 Modified exercise: Determining formulas for the lateral and surface areas of regular pyramids (Uma's Lessons Day 11).

Table 4.13 Frequency of Exercise Modifications Made by Uma

Original Exercise (any source)	Exercise as assigned by teacher	Code	Frequency
Reasoning-and-Proving Exercise	Exercise assigned, modified to LOWER RP	-1	0
Reasoning-and-Proving Exercise	Exercise assigned, neutral modification effect	0	1
Reasoning-and-Proving Exercise	Exercise assigned, modified to INCREASE RP	+1	0
Non-Reasoning-and-Proving Exercise	Exercise assigned, modified to INCLUDE RP	+2	2

4.2.1.4 Exercise sources *(RQ1d: What were the sources of the exercises that participants selected for instruction?)*

About two-thirds of the exercises selected by Uma were taken directly or modified from her textbook and accompanying workbook (see Table 4.14). Out of the exercises modified from the curriculum, three of them were reasoning-and-proving: make a conjecture from the area of a regular polygon lesson (#5 in Figure 4.14), make a conjecture and develop an argument for the formula for the surface area of a regular pyramid (see Figure 4.20 for both modifications), and investigate a conjecture from the chapter review exercises. The 61 exercises taken from ancillary resources were from two chapter exams written by Uma's department and given to all geometry students.

Table 4.14 Sources of Exercises Selected by Uma

Source	**All Exercises Frequency**	**Reasoning-and-Proving Exercises Frequency**
Taken Directly from Published Textbook/Curriculum	64	2
Modified from Textbook/Curriculum	129	3
Used in the CORP Course	0	0
Taken or Adapted from Ancillary Resources	61	0
Created by Teacher	36	5

Half of the reasoning-and-proving exercises Uma gave to her students were created by Uma herself. Two of these were exercises used in the first and third classroom artifact packets (investigate a claim/find a counterexample and make a conjecture/develop an argument, respectively) and three were for making conjectures and developing arguments for the areas of parallelograms, trapezoids, and rhombi. In her post-interview, Uma expressed a desire for a resource book of proof tasks. She did not identify any other sources of tasks that might currently exist, although her textbook contained 41 reasoning-and-proving exercises that Uma could have selected; either Uma was unaware that these exercises had the potential to engage her students in reasoning-and-proving or she felt too constrained by time to offer them to her students.

4.2.2 Implementation of reasoning-and-proving exercises *(RQ2: To what extent were participants able to maintain the level of cognitive demand of reasoning-and-proving tasks during implementation?)*

Of the three classroom artifact packets submitted by Uma, two were judged to contain a reasoning-and-proving task by the raters. These tasks were:

- "Prove or disprove. If two polygons have the same area, then they are congruent" (first classroom artifact packet, Day 3)
- "You have 36 feet of flexible fencing to build me the largest pen for all of my animals. You must determine what shape will give me the largest area, label the shape with its dimensions, and explain how you know that your shape has the largest area" (third classroom artifact packet, Day 8)

The task in the second student work sample involved a series of irregular shapes with scaffolded directions on how to dissect and calculate the area of the shapes (see Figure 4.21 for an example). Even though Uma asked her students to present their solutions and look for trends, the directions were so specific that there was little ambiguity about how to proceed (see the IQA rubric for Potential of the Task in Appendix K), so the task was rated by the coders to be a low-level task which did not involve reasoning-and-proof (scoring a 2 on Potential of the Task and a 2 on Implementation of the Task).

Directions: Calculate the Area of the Following Shape:

a. BREAK THE SHAPE UP into shapes that you know how to find the area of.
b. COLOR CODE the area of the sub-shapes. The same color should be used to WRITE THE FORMULA and FIND THE AREA of that shape.
c. EXPLAIN how you found the area of the irregular shape.

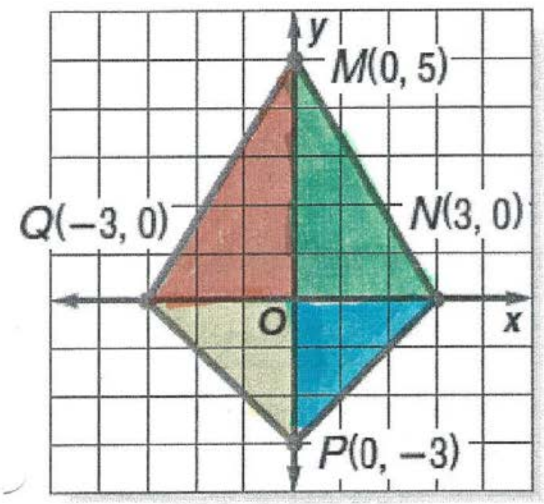


Figure 4.21 Uma's second classroom artifact packet task (Uma’s Lessons, Day 5).

Because the claim in Uma’s first classroom artifact packet (“If two polygons have the same area, then they are congruent”) is false, all that was required of the students was to find a counterexample. This still qualifies the potential of this task to be rated as a 4, since the IQA rubric states that “the task MAY require students to…find a counterexample” (see Appendix K). Student Sample Uma 1-1 first claims (correctly) that two figures with the same area can have different shapes (circle and square), then provides an example of two different quadrilaterals (parallelogram and rectangle) that have areas of 32 square units (see Figure 4.22). Student sample Uma 1-2 took a similar approach, approximating the square root of 12 as 3.5 to create a square and a triangle with the same area (see Figure 4.23). Student sample Uma 1-3 took a minimalist approach, and merely claimed that a square and a triangle can have the same area, but did not demonstrate this fact (see Figure 4.24). The last sample of student work that Uma judged to exceed or meet her expectations was Uma 1-4. This student's work indicates a lack of conviction or perhaps weak communication skills (see Figure 4.25). Since the task in this first classroom artifact packet asked students to prove or disprove a statement and all four samples that exceeded or met Uma's expectations disproved the statement to varying degrees, Uma maintained the high cognitive demand of this reasoning-and-proving exercise (scoring a 4 on the IQA rubric for Implementation of the Task).

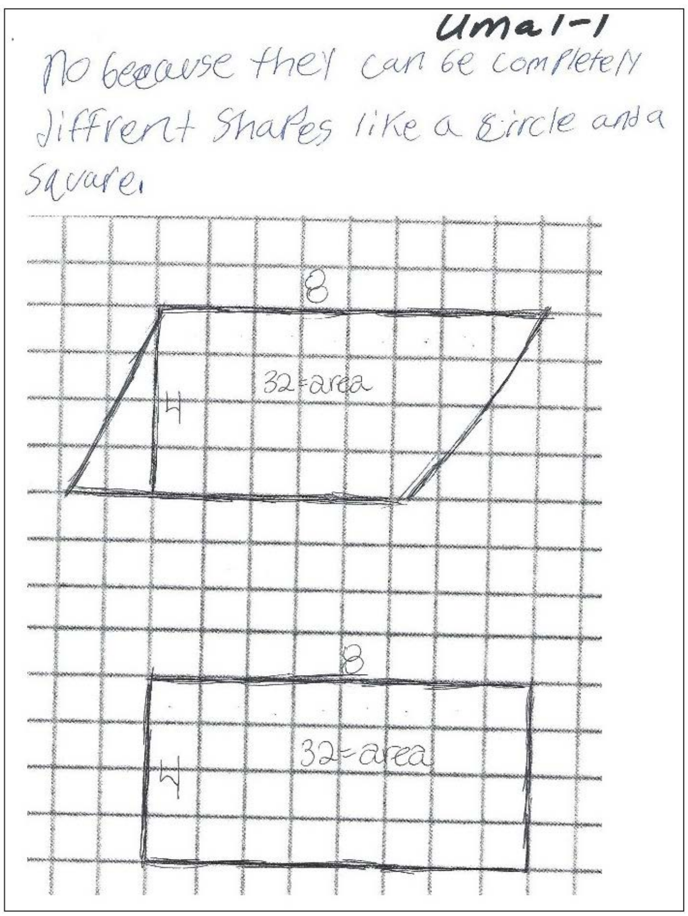


Figure 4.22 First classroom artifact packet sample Uma 1-1.

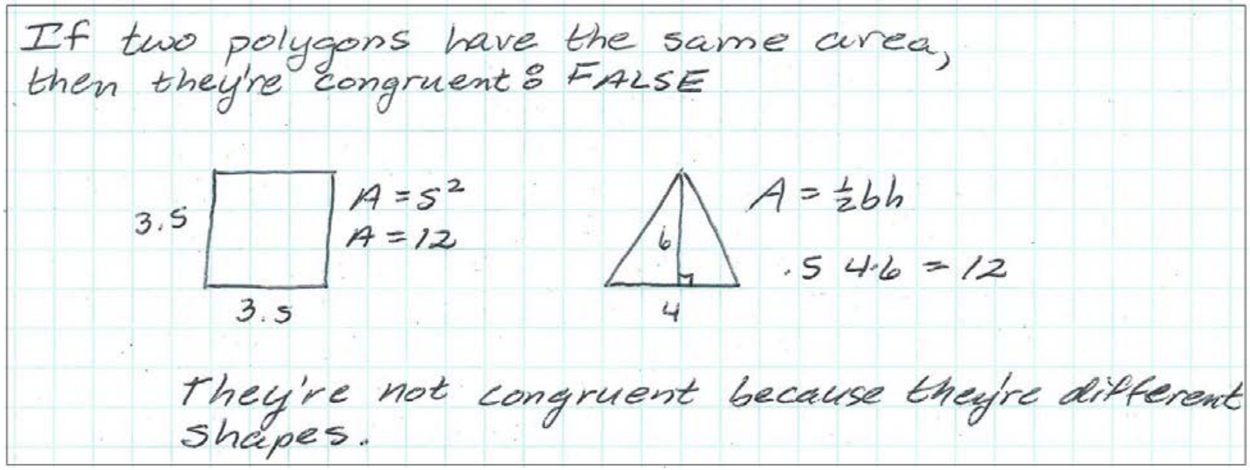


Figure 4.23 First classroom artifact packet sample Uma 1-2.

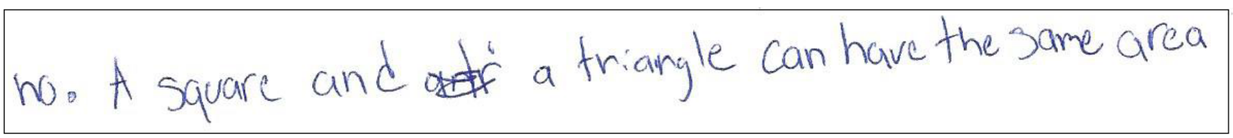


Figure 4.24 First classroom artifact packet sample Uma 1-3.

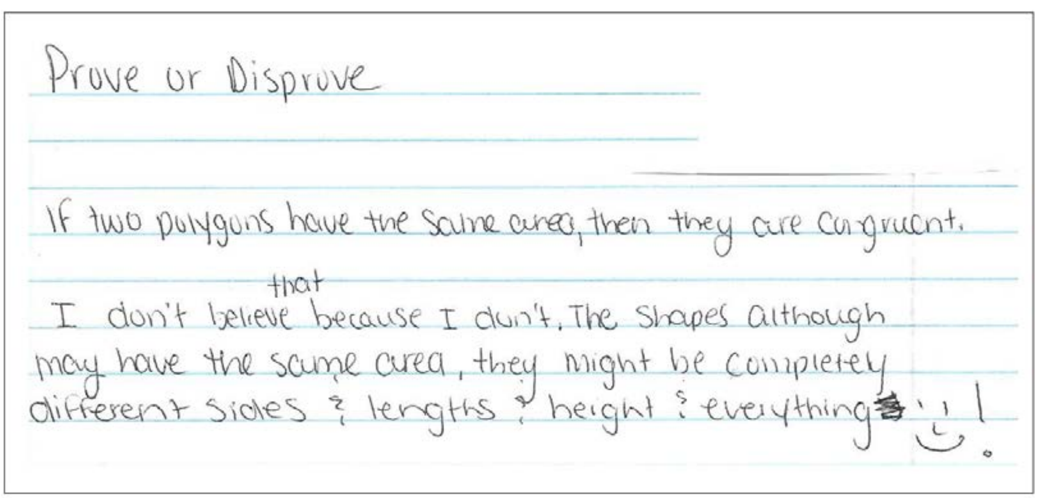
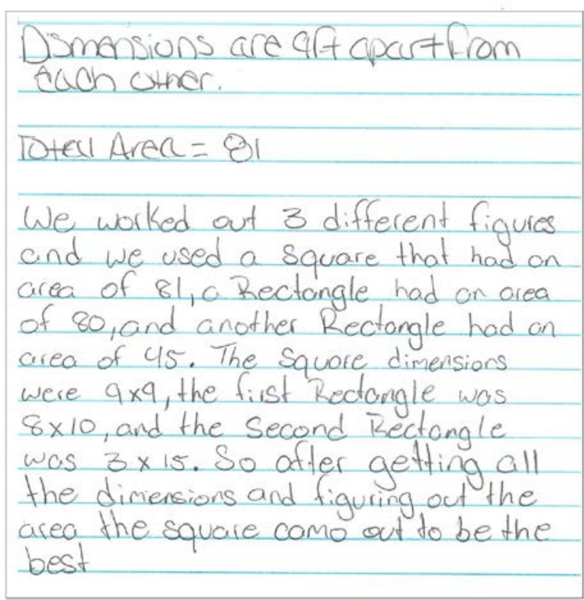


Figure 4.25 First classroom artifact packet sample Uma 1-4.

The third task required more of the students: they had to determine that a circle encloses more area than any polygon for a given circumference (perimeter) and they had to explain why. While most students calculated a series of areas of rectangles and squares (see Figure 4.2.13) and some students calculated the area of a circle (see Figure 4.2.14), only one student in the sample attempted to explain *why* a circle had the maximum area: “The circle is bigger because whatever shape you make it basically stretches the perimeter to the max because it has no corners” (sample Uma 3-2). The student then calculated some sample areas to reinforce the argument. Because this task required students to make explicit their reasoning but Uma did not hold students to determining and explaining why a circle encloses the maximum area for a fixed perimeter, this task was rated a 4 for task potential and 3 for task implementation (“students made conjectures but did not provide sufficient mathematical evidence or explanations to support conclusions”; see Appendix 3.9 for IQA Rubric 2: Implementation of the Task). A summary of these scores is shown in Table 4.2.5.


Dimensions are 9ft apart from each other.

Total Area = 81

We worked out 3 different figures and we used a square that had an area of 81, a Rectangle had an area of 80, and another Rectangle had an area of 45. The square dimensions were 9x9, the first Rectangle was 8x10, and the Second Rectangle was 3x15. So after getting all the dimensions and figuring out the area the square como out to be the best

Figure 4.26 Third classroom artifact packet sample Uma 3-3, showing an insufficient argument.

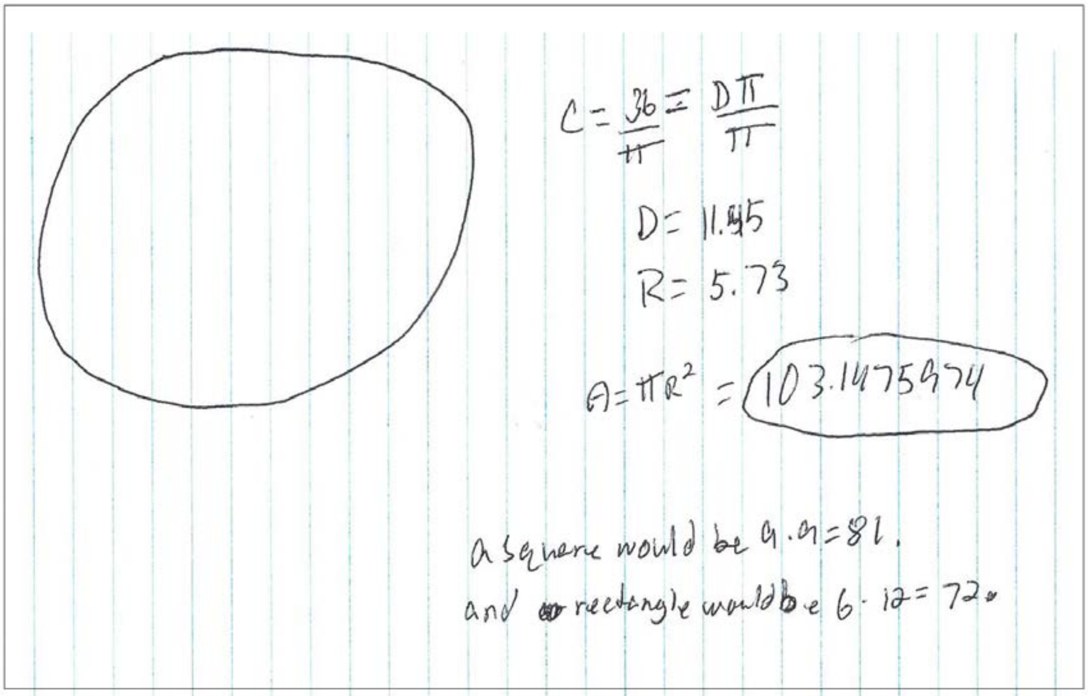


Figure 4.27 Third classroom artifact packet sample Uma 3-1 showing three calculations of area.

Table 4.15 Potential and Implementation of the Cognitive Demand of Uma's Tasks

Student Work Sample	Potential of Exercise	Implementation of Exercise	Maintain Cognitive Demand?
First	4	4	Maintain
Second	2	2	Maintain
Third	4	3	Maintain

4.2.3 Evaluation students' reasoning-and-proving products *(RQ3: To what extent were participants able to accurately evaluate their students' reasoning-and-proving products?)*

When asked to define reasoning-and-proving to a student and a colleague, Uma said she would give the same response: "A proof is a mathematical argument to illustrate some type of fact. You can use pictures, formulas, explanations in words to explain why your idea is true or not. It has to be something that will always work" (personal communication, May 29, 2013). Uma's definition loosely contains the first two core elements of proof (the argument must show that the conjecture or claim is—or is not—true for all cases and the statements and definitions that are used in the argument must be ones that are true and accepted by the community) but it does not contain any statement about the logic of the argument, which is another core element (see Table 4.16). Uma's definition did not mention type or form of proof, the representation used, or explanatory power, none of which are core elements and were correctly absent from Uma's definition of proof.

Table 4.16 Comparison of the Core Elements of Proof with Uma's definition of Proof

Core element of proof	**Uma**
The argument must show that the conjecture or claim is (or is not) true for *all* cases.	Present
The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.	Present
The conclusion that is reached from the set of statements must follow logically from the argument made.	Missing
Additional Criteria	None

4.2.3.1 Criteria for judgment *(RQ3a: To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?)*

Uma's rubrics for evaluating her students' work contained little information about what distinguished high, medium, and low performance. For instance, Uma provided the following scoring guide for the first student work packet task ("Prove or disprove that polygons with the same area are congruent"):

- A: picture (evidence), definition, neat and convincing
- B/C: correct answer with explanation, no pictures, not a good enough job convincing
- D/F: incorrect answer, no explanations

While this scheme is loosely based on the core elements of proof (disprove with a counterexample, correct formulas to calculate areas, decision is based on counterexample), it also contains a reliance on representation, which is not a core element of proof. The third student work sample contained a scoring guide that was vaguer than the first guide: "You will be graded on your correctness, explanations and diagrams, [and] behavior and participation in the conversation that compares each group's work." Uma's rubrics are summarized in Table 4.17. It was assumed by the raters that Uma considered "correctness" in the third rubric to mean that the students correctly determined that a circle would enclose the largest area. It should also be noted that students' behavior, neatness and presentations factored into the clarity and detail of Uma's expectations. Because the coder was not present in the classroom during these presentations, the remotely-gathered data may over- or under-represent the standard to which Uma held her students with respect to making their reasoning explicit during the presentations. The core elements of proof should have been present in the rubrics for the first and third exercise (the first exercise directs the students to "prove or disprove" and the third exercise directs student to "explain how you know that your shape has the largest area"). The presence of the core elements of proof for a proof task is a requirement to earn a 4 on the IQA rubric Clarity and Detail of Expectations (see Appendix K). In her post-interview, Uma stated that the class struggled in the third exercise with determining a circle to be the correct shape which inspired a lengthy discussion (personal communication, May 29, 2013), The IQA Clarity and Detail of Expectations rubric scores are in Table 4.18.

Table 4.17 Uma's Rubrics for Her Classroom Artifact Packets

First Rubric	Second Rubric	Third Rubric
A: Picture (evidence), definition, neat, convincing B/C: Correct answer with explanations, no pictures, not a good enough job convincing D/F: Incorrect answer, no explanations	Correctness, neatness, presentation, behavior, and answers to questions on the back (Note: all groups met expectations)	Correctness, explanation and diagrams, behavior and participation in the conversation that compares each group's work

Table 4.18 Quality of Uma's Rubrics Used to Judge Student Work

Classroom Artifact Packet Rubric	Clarity and Detail of Expectations	Comment
First	3	Vague differentiation between levels
Second	3	Vague differentiation between levels
Third	2	Broadly stated and unelaborated
Average	*2.7*	

4.2.3.2 Application of core elements of proof ***(RQ3b: To what extent did participants apply the core elements of proof in evaluating their students' reasoning-and-proving products?)***

Uma's expectation scores for her students' work somewhat aligned with her evaluation of whether or not the work qualified as a proof, disproof, or non-proof argument. In one case (Uma 1-3; see Table 4.19), Uma's evaluation of her students' disproof attempt was not the same as the primary researcher's evaluation. The primary researcher thought that stating that a "square and a triangle can have the same area" was sufficient evidence to disprove the claim that if two polygons have the same area, then the polygons are congruent. Uma's evaluation required that students provide pictures as evidence. It may be that Uma—with her knowledge of her students' abilities and requirements for evidence—thought that the only way students could convincingly justify their decision was to provide contrasting examples (burden of proof depends on the classroom community who is the audience for the argument). On the other hand, representation is not one of the core elements of proof.

Table 4.19 Uma's Application of the Core Elements of Proof in Evaluating Student Work

Classroom Artifact Packet Sample	Participant Score	Participant Evaluation: Proof?	Researcher Evaluation: Proof?	Comments
First task: Prove or disprove. If two polygons have the same area, then the polygons are congruent.				
Uma 1-1 (Figure 4.22)	Exceeded expectations	Disproof	Disproof	*Drew a parallelogram and a rectangle with an area of 32 units^2*
Uma 1-2 (Figure 4.23)	Exceeded expectations	Disproof	Disproof	*Drew a rectangle and a triangle with an area of 12 units^2*
Uma 1-3 (Figure 4.24)	Met expectations	Nonproof	Disproof	*"No. A square and a triangle can have the same area." (no pictures or examples)*
Uma 1-4 (Figure 4.25)	Met expectations	Nonproof	Nonproof	*"I don't believe that because I don't. The shapes although may have the same area, they might be completely different sides & lengths & height & everything!" (no pictures or examples…no "polygons")*
Uma 1-5	Failed expectations	Nonproof	Nonproof	*Student drew two 5x5 squares and two irregular hexagons of different sizes.*
Uma 1-6	Failed expectations	Nonproof	Nonproof	*"This is true if they have the same side and angles then they are congruent because all the sides will be equal and the angles will be the same also so yes this is very true. Even if they don't look like the pictures above [two 3 x 4 rectangles], they could look like this [4 x 3 rectangle and a 3 x 4 rectangle].*

Second task: Calculate the areas of irregular shapes

Uma 2-1	Met expectations		n/a	*Not a proof task*
Uma 2-2	Met expectations		n/a	
Uma 2-3	Met expectations		n/a	
Uma 2-4	Met expectations		n/a	
Uma 2-5	Met expectations		n/a	
Uma 2-6	Met expectations		n/a	

Third task: You have 36 feet of flexible fencing to build me the largest pen for all of my animals. You must determine what shape will give me the largest area, label the shape with its dimensions, and explain how you know that your shape has the largest area.

Uma 3-1 (Figure 4.27)	Exceeded expectations	Nonproof	Nonproof	*Empirical (calculated a circle, a rectangle, and a square area)*
Uma 3-2	Exceeded expectations	Proof	Proof	*Provided reason for why circle contains the maximum area*
Uma 3-3 (Figure 4.26)	Met expectations	Nonproof	Nonproof	*Empirical (calculated a square and two rectangle areas)*
Uma 3-4	Met expectations	Nonproof	Nonproof	*Empirical (calculated 10 rectangles and one square)*
Uma 3-5	Failed expectations	Nonproof	Nonproof	*"irregular is bigger than regular (octagon)"*
Uma 3-6	Failed expectations	Nonproof	Nonproof	*Drew a "T" shaped area and counted boxes for area*

4.2.3.3 Communication of expectations *(RQ3c: In what ways did participants communicate expectations regarding what is required to produce a proof to students?)*

Uma's written descriptions of how she communicated her expectations to her students were a little vague, as reported on the classroom artifact packet cover sheets. In her post-interview, Uma clarified that for each classroom artifact packet task, she did discuss (but did not model) her expectations for her students in advance of their completing the assignment. These results and comments are in Table 4.20.

Table 4.20 Uma's Communication of Expectations

Classroom Artifact Packet	Communication Score	Comments
First	3	Printed on task: "You may use pictures, words, formulas to help you explain your answer. You may use graph paper, high-lighters, rulers etc. to help you explain." Verbal instructions: "Must first work alone then collaborate with group. Group answer to be graded."
Second	3	Printed on task: "Break up the shapes…color code the areas of sub-shapes…explain" (see Figure 4.2.8). Verbal instructions: "With group must determine area, then present to class. As they are presenting, other groups take notes, ask questions. Graded on work and behavior."
Third	3	Printed on task: "You must determine what shape will give me the largest area, label the shape with its dimensions, [and] explain how you know that your shape has the largest area. You will be graded on your correctness, explanation and diagrams, behavior and participation in the conversation that compares each group's work."
Average Score	3	

The geometry class that Uma selected for this study is a daily double-block class, designed for struggling students. Uma reported that she had the lowest-level students in the school; many had already failed geometry once but needed to pass geometry to earn their diploma (personal communication, February 24, 2013). On her first classroom artifact packet cover sheet, Uma wrote that her students typically did one proof task per chapter, "but it is not a skill they practice often." For the task of the third classroom artifact packet, Uma wrote that the pen task was "much more open ended than most tasks I give." Uma reported that she was "still working on trying to get my students to work on high-level tasks in groups" but since she was not afforded much time due to her districts strict pacing guide, her students did not have a chance to routinize their behavior in groups (personal communication, February 24, 2013 and May 29, 2013). This could explain why much of her communication of expectations to students involved behavior as opposed to core elements of proof.

In summary, Uma offered 10 reasoning-and-proving exercises to her students over the course of 15 instructional days during which students studied the area of two-dimensional shapes and the surface area of three-dimensional shapes. While 10 exercises represents a low percentage of exercises, it does represent about 16% of the time students spent in class. The textbook contained 41 reasoning-and-proving exercises (only 5 of which were selected by Uma). Most of the available and selected exercises asked students to make or investigate a conjecture, with a smaller number of exercises asking students to develop an argument. Uma modified 133 exercises but largely to include scaffolding rather than to increase the exercises' reasoning-and-proving potential. In two cases, however, Uma modified a formula presentation in the textbook into a discovery lesson which created opportunities to engage her students in reasoning-and-proving. In another exercise, Uma asked students to investigate a conjecture but did not label this particular exercise as reasoning-and-proving (she had labeled the entire chapter review game as non-reasoning-and-proving).

Two of the student work sample exercises had the potential to engage students in cognitively challenging reasoning-and-proving work; Uma maintained the cognitive demands during implementation. In evaluating her student's reasoning-and-proving products, Uma did not specifically refer to the core elements of proof in her rubrics but her ratings on her student work with respect to proof, disproof, and non-proof arguments generally agreed with the primary researcher and second coder ratings. Finally, Uma shared and discussed her expectations for work with her students prior to their work on each exercise in the student work packets.

4.3 SIDNEY

Sidney teaches in a middle school in a district just outside of a major metropolitan area in Virginia. His school serves 1200 students in grades 7 and 8 and is situated in an affluent area with academically-inclined parents. Many of his students test into a nationally-ranked high school known for its science and technology program. Unlike Uma's school, the Virginia Standards of Learning represent an easy bar for Sidney's students to pass. Sidney's school uses Prentice Hall's *Algebra 1* textbook, Virginia edition (Charles, Hall, Kennedy, Bellman, Bragg, Handlin, Murphy, & Wiggins, 2012), but each student is not assigned a textbook. Therefore, Sidney's department creates and shares lesson note sheets and assignments which are distributed to the students. Sidney can also generate worksheets with an online database from Prentice Hall or from Kuta software.

This was Sidney's first year teaching. He was assigned Math 7 and Algebra I; he chose Algebra 1 for the current study. Sidney was encouraged by the other members of his department to use the shared repository of notes and assessments to prevent "backlash" (personal communication, February 26, 2013). Sidney reported that while his district's pacing guide does not include much reasoning-and-proving, the members of his department and district realize the need for more reasoning-and-proving in their curriculum. During one in-service day this winter, Sidney was assigned to an all-day session with other math teachers from across the district to discuss reasoning-and-proving. Sidney says the intention of his department this summer is to combine less complicated learning standards (which are currently taught one per day) in order to free up more time for reasoning-and-proving.

4.3.1 Selecting exercises *(RQ1: To what extent did participants select reasoning-and-proving learning opportunities in the form of exercises?)*

4.3.1.1 Available in textbook *(RQ1a: To what extent does the textbook include exercises that have the potential to engage students in reasoning-and-proving?)*

Sidney's 15-day data collection spanned a unit on polynomials. The textbook offered 643 exercises (practice and problem solving, lesson check, mixed review, and standardized test prep) in the sections covered by Sidney's district standards; it was not possible to count and code the exercises available in the online or Kuta software exercise generators. 33 of the 643 exercises were reasoning-and-proving exercises. This represents 5.1% of the exercises. Students had the opportunity to make 14 conjectures, investigate 6 statements, correct 9 mistakes, develop 7 arguments, and find 1 counterexample in the 643 exercises (see Table 4.21).

Table 4.21 Available Types of Reasoning-and-Proving Exercises in Sidney's Textbook

Type of RP Exercise	Available in Textbook*
Make a Conjecture	14 (2.2%)
Investigate a Conjecture	6 (0.9%)
Evaluate an Argument	0 (0%)
Correct a Mistake	9 (1.4%)
Develop an Argument	7 (1.1%)
Counterexample	1 (0.2%)
Principles of Proof	0 (0%)
Non-Reasoning-and-Proving	610 (94.5%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

4.3.1.2 Selected by Sidney (RQ1b: To what extent did the participant select exercises for instruction that had the potential to engage students in reasoning-and-proving?)

Over the course of the 15 days, Sidney engaged his students in 657 exercises. These included 227 exercises done during instruction, 402 homework exercises, and 28 assessment exercises. Any exercise generated by the online textbook resource or Kuta software and was used by Sidney was counted as selected by Sidney. It should be noted that the homework worksheets offered by Sidney to his students contained between 17 and 72 exercises each and his students' homework is not collected and graded (thus making 61.1% of the exercises Sidney selected optional). Of the 657 exercises that Sidney selected, 14 were reasoning-and-proving exercises

(2.1%). On Day 1, Sidney opened his lesson on classifying, adding and subtracting polynomials (20 exercises in 23 minutes) with a make a conjecture/develop an argument exercise about the cost of *m* people in *n* cars visiting a museum and/or an aquarium. On Days 8/9 (17 exercises in 92 minutes), Sidney asked his students to "come up with some type of conjecture that explains how to divide with [sic] these kinds of polynomials" (binomials and trinomials by a monomial). On Day 11 (9 exercises in 23 minutes), Sidney asked his students to "write a conjecture that describes [how to factor a quadratic trinomial]." On Day 13 (18 exercises in 23 minutes), Sidney asked his students to write conjectures on procedures to factor perfect-square trinomials and the difference of two squares, then he asked his students to test his conjectures on four examples (investigate a conjecture). In addition, the optional homework contained 5 reasoning-and-proving exercises (4 correct a mistake and one develop an argument). In total, Sidney selected 5 make a conjecture, 4 investigate a conjecture, 2 develop an argument, and 4 correct a mistake exercises (see Table 4.22 and Figure 4.23); one of these exercises was double-coded because it asked students to make a conjecture *and* develop an argument.

Table 4.22 Selected Types of Reasoning-and-proving Exercises by Sidney

Type of RP Exercise	Selected by Sidney*
Make a Conjecture	5 (0.8%)
Investigate a Conjecture	4 (0.6%)
Evaluate an Argument	0 (0%)
Correct a Mistake	4 (0.6%)
Develop an Argument	2 (0.3%)
Counterexample	0 (0%)
Principles of Proof	0 (0%)
Non-Reasoning-and-Proving	643 (97.9%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

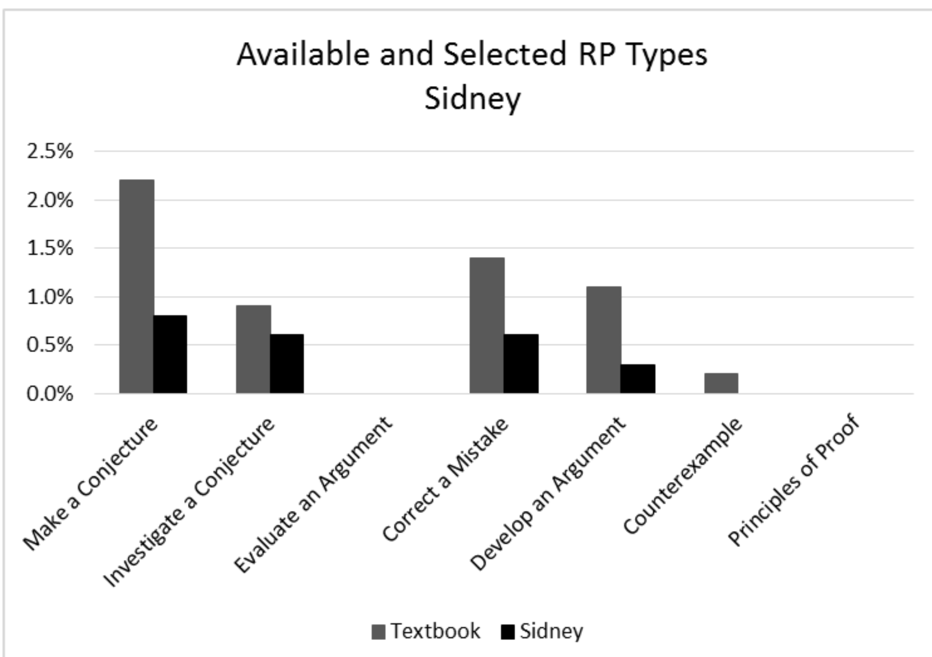


Figure 4.28 Comparison of Available and Selected Types of Reasoning-and-proving Exercises for Sidney

Of the 14 reasoning-and-proving exercises selected by Sidney for his lessons, Sidney labeled 10 of them as reasoning-and-proving and did not label 4 of them (the four investigate a conjecture exercises from Day 13). Of the 643 selected lesson exercises rated as non-reasoning-and-proving by the coders, Sidney labeled 19 of them as reasoning-and-proving and 624 as not reasoning-and-proving. Like Uma, however, Sidney collectively labeled large groups of exercises as either reasoning-and-proving or not on his Task Log Sheets. Fortunately, the primary researcher was able to meet with Sidney to get a more accurate picture of how Sidney viewed each exercise. The 19 mislabeled exercises represent exercises Sidney believed to have the potential to engage his students in reasoning-and-proving. Figure 4.29 is an example of one such exercise. The exercise asks students to explain their reasoning but it is not the type of reasoning that leads to proof. Figure 4.30 is another example that shows a set of related exercises (from Sidney's Lessons Day 13) but only the last exercise (part d) asks for a conjecture; Sidney labeled the entire set as having the potential to engage students in reasoning-and-proving.

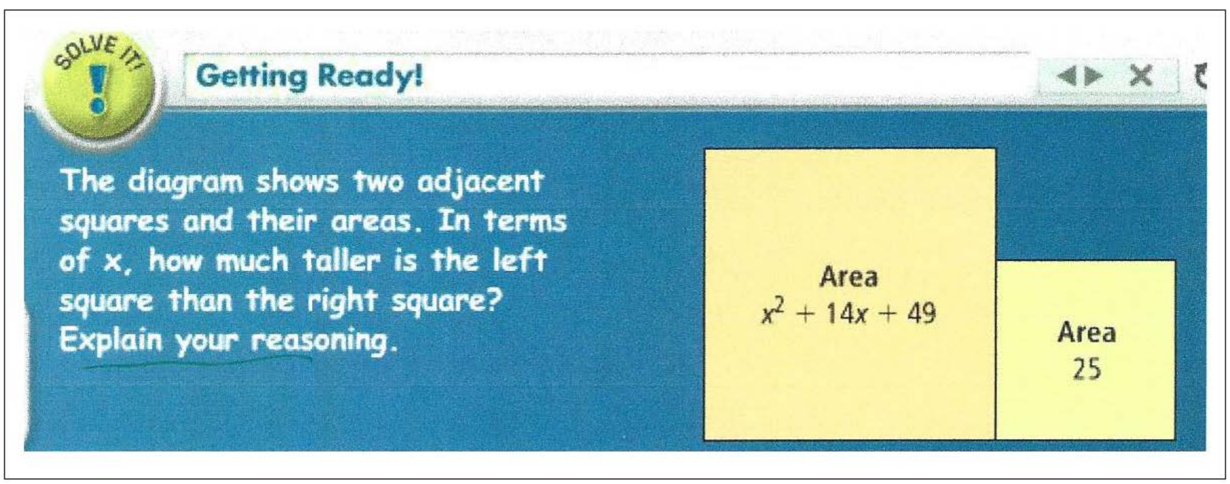


Figure 4.29 Mislabeled exercise from Sidney's lessons Day 13 (Charles et al., 2012a, p. 511).

Factoring Perfect-Square Trinomials:

a. Recall how we can multiply perfect square trinomials from a previous lesson:
b. Using the example, $x^2 + 8x + 16$, label the parts that corresponds [sic] to part a?
c. Factor the example back into the two binomials?
d. Write a conjectures [sic] on how we can accopmlish [sic] this every time.

Figure 4.30 Collection of exercises Sidney labeled as "reasoning-and-proving."

4.3.1.3 Exercise modifications *(RQ1c: To what extent did the participant modify exercises to affect the exercises' potential to engage students in reasoning-and-proving?)*

In the 15-day data collection period, Sidney modified 19 tasks. Three of the modifications involved reasoning-and-proving exercises. In these three modifications, Sidney asked his students to make conjectures about procedures. On Day 11, Sidney asked his students to represent the quadratic expression $x^2 + 4x + 4$ with algebra tiles. He then wrote,

> Now, your goal is to work backwards to get the two binomials that form this trinomial. Go ahead and using the picture you just drew, write the two binomials along the sides of the model. Write a conjecture that describes about [sic] how to return back to the original two binomials (Sidney' Lessons Day 11).

Sidney's answer key had the target response: "You want the sum to equal the middle term and the product of these numbers equal constant [sic]." The other two modifications regarding conjectures occurred on Day 13, when students were studying factoring special cases of quadratic expressions (e.g., difference of squares). Sidney added, "write a conjectures [sic] on how we can accopmlish [sic] this every time" to the notes on factoring perfect square trinomials (see Figure 4.30). Sidney made a similar modification in his lesson notes regarding factoring a difference of two squares.

Sidney's modifications are summarized in Table 4.23. None of the remaining 16 modifications were on reasoning-and-proving problems so they are not represented in the table, however, the modifications are described here. Of these 16, six modifications to exercises occurred because Sidney lost a day of instruction due to standardized testing and he created a video lesson to replace instructional time. Another 3 modifications provided scaffolding to his students (e.g., "record two key aspects of…", "review the process of long division…", and "write down everything you know about…"). Sidney also switched the numbers or signs in three exercises in the department notes, asked students to model factoring with Algebra Tiles twice, and asked students to choose or invent a method to accomplish a goal twice.

Table 4.23 Frequency of Exercise Modifications Made by Sidney

Original Exercise (any source)	Exercise as assigned by teacher	Code	Frequency
Reasoning-and-Proving Exercise	Exercise assigned, modified to LOWER RP	-1	0
Reasoning-and-Proving Exercise	Exercise assigned, neutral effect of modification	0	0
Reasoning-and-Proving Exercise	Exercise assigned, modified to INCREASE RP	+1	0
Non-Reasoning-and-Proving Exercise	Exercise assigned, modified to INCLUDE RP	+2	3

4.3.1.4 Exercise sources (*RQ1d: What were the sources of the exercises that participants selected for instruction?)*

Sidney took over half of his exercises from his textbook or the online exercise generator, about a fifth from Kuta software, and about a sixth from other teachers (which included lesson exercises, the mid-chapter assessment, and homework exercises). Specifically, the "taken directly from published textbook/curriculum" exercises consisted of printed textbook problems and those selected by Sidney from the publisher's problem generator, and the ancillary resource exercises consisted of 105 exercises from other teachers and 135 exercises selected by Sidney from Kuta software. Sidney based his lessons on the lessons found in the shared departmental repository; many of the exercises in the other teachers' lessons came from the published textbook. Sidney used the online teacher resource center exercise generator to create worksheets used for homework (338 exercises, about half of all of the exercises), so those problems were selected by Sidney but not created by Sidney. The data is summarized in Table 4.24.

Table 4.24 Sources of Exercises Selected by Sidney

Source	All Exercises Frequency	Reasoning-and-Proving Exercises Frequency
Taken Directly from Published Textbook/Curriculum	390 (59.3%)	10
Modified from Textbook/Curriculum	7 (1.1%)	2
Used in the CORP Course	0	0
Taken or Adapted from Ancillary Resources	240 (36.5%)	0
Created by Teacher	20 (3.0%)	2

4.3.2 Implementation of reasoning-and-proving exercises *(RQ2: To what extent were participants able to maintain the level of cognitive demand of reasoning-and-proving tasks during implementation?)*

Sidney's 15-day data collection period covered the mathematical content polynomials: definitions and classification, operations, factoring, and simplifying a function with radicals in its denominator by rationalization. Sidney's first classroom artifact packet task is shown in Figure 4.31.

> 10.3 Multiply a polynomial by a monomial
>
> Consider the following example: Solve: $2x(3x + 1)$
>
> Write down everything you know about this expression:
>
> What is your solution? Show your work!!
>
> Explain how you get that solution using information you already know.
>
> See if your method works for this expression: $-x^3(9x^4 - 2x^3 + 7)$
>
> Now, try solving it another way different then [sic] the way you solved it the first time.

Figure 4.31 Sidney's first classroom artifact packet task (Day 2).

In this task (the collection of exercises shown in Figure 4.31), students were asked to determine a method for multiplying a polynomial by a monomial (the hope was that students would apply the Distributive Property). Students had already studied the distributive property and simplifying exponents with the same base earlier in Algebra 1 and in a previous course

(Math 7), and most of the students solved the first exercise in the same way. Thus, while this task could have had the potential to engage some students in cognitively demanding work, it was not demanding for most of the students. In the four samples of student work Sidney submitted, his students explained their method in these ways:

- "multiply everything by 2x = distributive property" (Sidney 1-1, exceeded expectations)
- "I know the distributive property works" (Sidney 1-2, met expectations)
- No response (Sidney 1-3, failed expectations)
- "I had to multiply the linear monomial by each term of the linear binomial, through the distributive property" (Sidney 1-4, failed expectations)

Interestingly, the student labeled Sidney 1-4 offered the most complete and clear explanation of her process, but because she did not finish the worksheet, she received a "failed expectations" evaluation. Since the students' work suggests that "there is little ambiguity about what needs to be done and how to do it" (IQA Rubric for Potential of the Task; see Appendix K), it was rated a 2 for potential and implementation by the coders.

In Sidney's second student work sample task (see Figure 4.32), students are asked to determine how to write the lengths of the reduced piece of paper (6 – x) and then find an expression for the area of the resulting invitation. This task was a little more complicated than the first student work sample task because students had to create an expression for the new lengths and then substitute their expression into the formula for the area of a square. In Sidney's notes from Days 2 and 3 (combined), students worked on a similar problem involving the extension of a patio door. Because the students had recently seen a similar problem and had already practiced multiplying binomials, the task potential was not rated as a 4 ("there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task"; IQA Rubric Potential of the Task). Neither is it a 2 ("there is little ambiguity about what needs to be done and how to do it"; IQA Rubric Potential of the Task; see Appendix K) because the students still had to write an expression for the new sides of the invitation. Thus, the coders rated the potential of this task as 3. The similarity in the students' answers, however, suggested that "students engaged in using a procedure [whose]…use was evident based on prior instruction, experience, or placement of the task" (IQA Rubric Implementation of the Task level 2; see Appendix K). Notice that the task does not suggest a form of the expression for the area; a student reporting $A = (6 - x)^2$ would have been correct. The lesson for this day was, however, about multiplying special cases and most of the student work samples (see Figures 4.33 and 4.34) show that students did convert their expressions for area into standard form by multiplying the expression for side length $(6 - x)$ by itself.

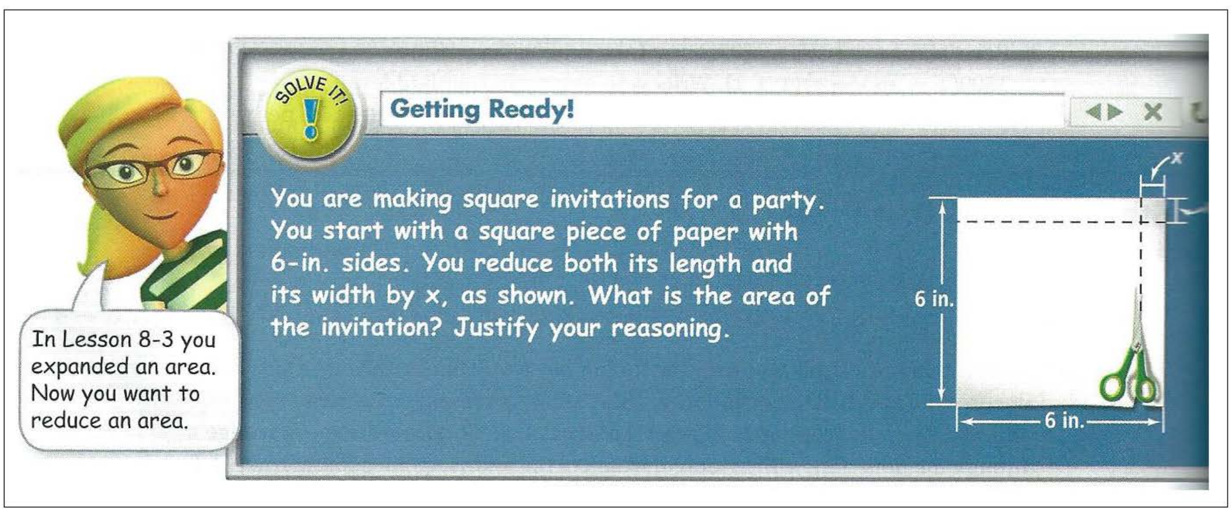


Figure 4.32 Sidney's second classroom artifact packet task (Day 5) (Charles et al., 2012a, p. 492).

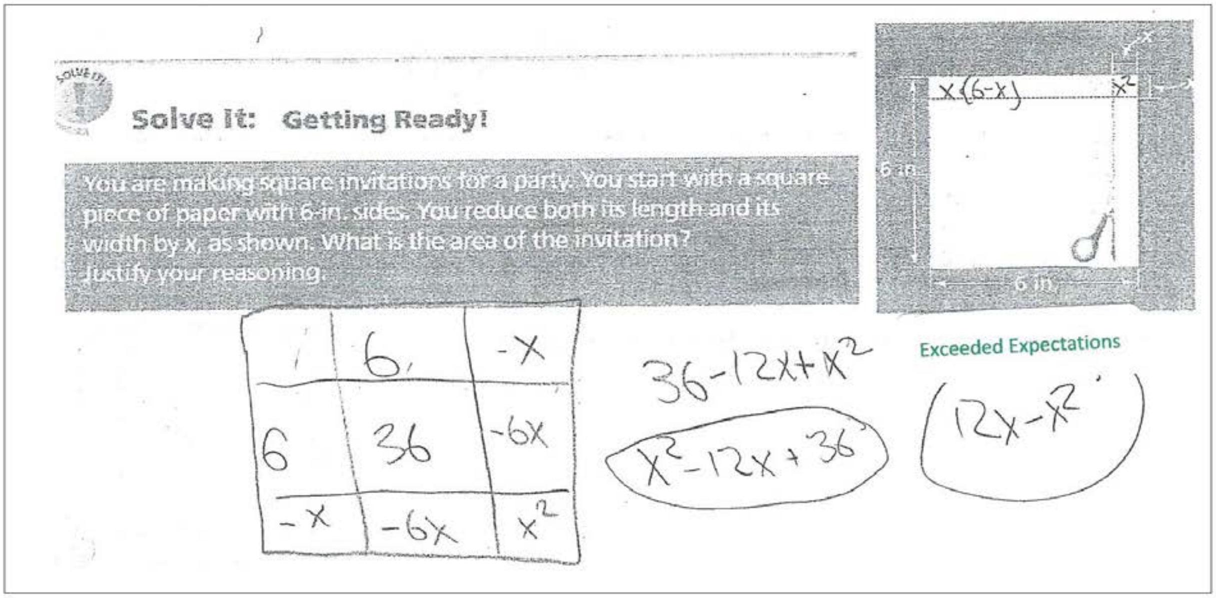


Figure 4.33 Second classroom artifact packet sample Sidney 2-1.

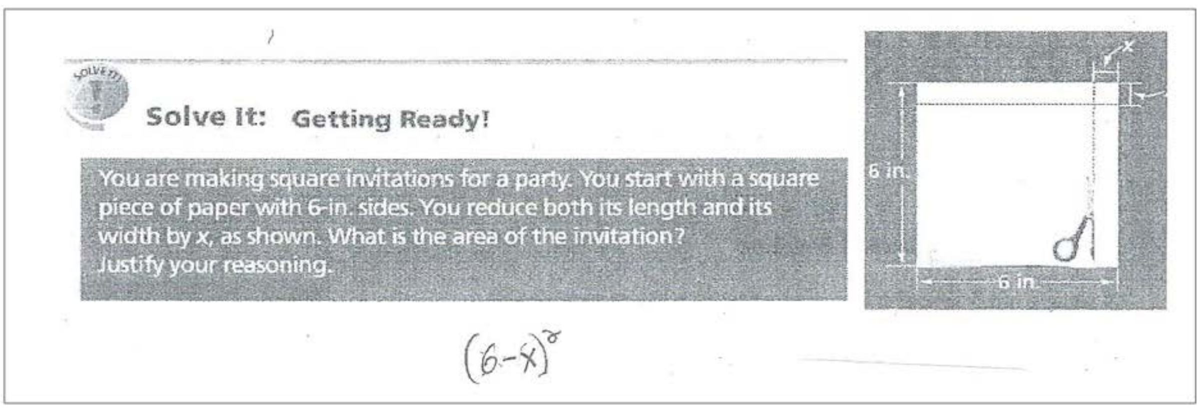
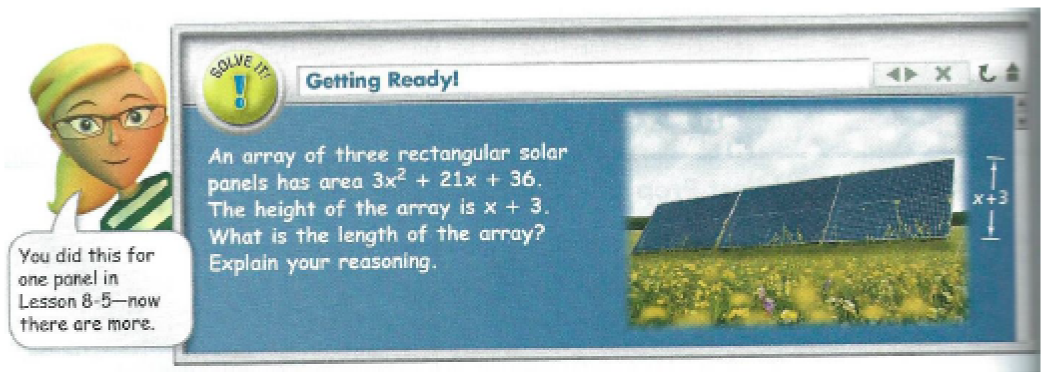


Figure 4.34 Second classroom artifact sample Sidney 2-5.

In the middle of his third student work sample task (see Figure 4.35), Sidney asked students to write a conjecture that described how to factor a trinomial with a leading coefficient of 1. While this is a conjecture, it is not a conjecture that would lead to a proof. The task is about creating a procedure or algorithm and thus is not considered a reasoning-and-proving task.

Since the task prompts students to use algebra tiles (which students had practiced using for multiplying binomials), the "task does not explicitly prompt for evidence of students' reasoning and understanding" (IQA Rubric Potential of the Task, score 3; see Appendix K). For these reasons, the potential of the task was rated as a 3 by the coders.

10.9 Factoring quadratic trinomials



General Form of a trinomial:

With the algebra tiles, demonstrate the following expression: $x^2 + 4x + 4$

Draw your expression here:

Now, your goal is to work backwards to get the two binomials the form this trinomial. Go ahead and use the picture you just drew, write the two binomials along the side of the model.

Write a conjecture that describes about how to return back to the original two binomials.

What did we come up with?

Two key aspects when we ____[factor]_____ a trinomial?

Figure 4.35 Sidney's third classroom artifact packet task (Day 11) (adapted from Charles et al., 2012a, p. 506).

Sidney's students wrote the following conjectures for this procedure:

- "I am looking for a number when you multiply them you get the constant (c) and when you add them you get the middle term (b)" (Sidney 3-1)

- "You look for numbers that when you multiply them together, you get the constant and when you add them you get the middle term (b)" (Sidney 3-2)
- "I am looking for numbers that, when multiplied together, I get the constant (c), and when I add them, I get the middle-term coefficient (b)" (Sidney 3-3)
- "I am looking for numbers, when I multiply them, I get the constant (c), and when I add them, I get the middle term (b)." (Sidney 3-4)
- No conjectures recorded (Sidney 3-5 and Sidney 3-6)

These conjectures are remarkably similar in their wording and notation, especially after only two examples (lesson Getting Ready! Problem: divide $x^2 + 7x + 12$ by $x + 3$ and demonstrate with algebra tiles the expression $x^2 + 4x + 4$; Sidney Lesson Day 11). This suggests that students had already seen this procedure prior to this lesson or it was discussed in class before the conjectures were recorded. Table 4.25 summarizes the potential and implementation of each of Sidney's student work sample tasks.

Table 4.25 Potential and Implementation of Cognitive Demand of Sidney's Tasks

Classroom Artifact Packet	Potential of Exercise	Implementation of Exercise	Maintain Cognitive Demand?
First	2	2	Maintained
Second	3	2	Declined
Third	3	2	Declined

4.3.3 Evaluation students' reasoning-and-proving products *(RQ3: To what extent were participants able to evaluate their students' reasoning-and-proving products?)*

When asked how Sidney would define proof to a colleague or student, Sidney replied, "Proof is when you are able to show beyond a reasonable doubt that something is true, using previous knowledge that is already known. Proof is not just using examples showing something works, but why it works" (personal communication, May 26, 2013). In comparing Sidney's definition with the core elements of proof, Sidney's definition came close to the first core element of proof (that the argument must show that the conjecture or claim is—or is not—true for *all* cases), but left out "all cases." His definition also included statements and definitions that are used in the argument must be ones that are true and accepted by the community when he included "using previous knowledge that is already known." He did not include anything about a logical flow to the argument, and he included explanatory power, which is not a requirement of a valid proof (see Table 4.26).

Table 4.26 Comparison of the Core Elements of Proof with Sidney's Definition of Proof

Core element of proof	**Sidney**
The argument must show that the conjecture or claim is (or is not) true for *all* cases.	Partial
The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.	Present
The conclusion that is reached from the set of statements must follow logically from the argument made.	Missing
Additional Criteria	partial

Since none of Sidney's classroom artifact packet tasks asked students to construct a proof, none of the samples of student work were evaluated with respect to whether or not the students' argument counted as proof. It should be noted that Sidney confirmed that he intended these tasks to be proof tasks, but his students did not give him "the result that [he] was looking for" (personal communication, July 5, 2013). In the CORP course, Sidney and his fellow participants spent time creating and comparing different types of proofs (visual, algebraic, etc.). Many of the tasks for which participants were asked to create a proof were pattern tasks. For a pattern task, participants generated data and drew figures, then made a conjecture which was a generalized statement about the observed pattern (e.g., a formula). For instance, the first proof task in which the participants engaged was the "Squares Problem," where participants were asked to determine how many different 3-by-3 squares there are in a 60-by-60 square and why their answer was correct. Most participants created formulas (conjectures) to determine the answer; the course instructor then led a rich discussion which helped participants link elements of their formulas with the context of the problem and engaged participants in comparing and contrasting different versions of the formula. It was through this discussion that participants *persuaded* each other that their answers were correct (thus, *proving* their conjecture would hold true for a N-by-N square). Thus, it is possible that when Sidney asked his students to create "multiple methods" and model with algebra tiles he was trying to help them ascertain and persuade as was done in the CORP course.

4.3.3.1 Criteria for judgment *(RQ3a: To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?)* Sidney's criteria for judging his students' work products "was more general" (third classroom artifact packet cover sheet) and included the following elements:

1. Where [sic] they able to reach the goal?
2. Did they use the algebra tiles or just drawings?
3. Did they do something completely on there [sic] own?
4. Did they explain why or just provide answers?
5. Were they able to come up with a conjecture and use it?

Sidney said that he used the same criteria for all three classroom artifact packets. When pressed for differentiation between exceeded, met, and failed expectations, Sidney provided the following information on levels he used for each sample of student work (personal communication, May 26, 2013):

- Low: didn't finish, errors, misconceptions
- Medium: did not exceed; did only what was necessary
- High: used multiple ways, no errors or fixed errors

Initially, the primary researcher assigned Sidney a Clarity and Detail of Expectations rubric score of 2, but after asking Sidney for more details about his distinctions between low, medium, and high scores, the coders raised Sidney's Clarity and Detail of Expectations score to a 3 (see Table 4.27).

Table 4.27 Quality of Sidney's Rubrics Used to Judge Student Work

Classroom Artifact Packet Rubric	Clarity and Detail of Expectations	Comment
First	3	"Able to explain thoroughly, their conjecture worked, use previous knowledge"
Second	3	"Did students try another way, was their reasoning sound? Did the [sic] pull past knowledge or ask a neighbor? Did they just wait and copy it down"
Third	3	"Where [sic] they able to reach the goal, did they use the algebra tiles or just drawings, they do something completely on there [sic] own, did they explain why or just answers, where [sic] they able to come up with a conjecture and use it?"
Average	*3*	

4.3.3.2 Application of the core elements of proof *(RQ3b: To what extent did participants apply the core elements of proof in evaluating their students' reasoning-and-proving products?)*

As previously mentioned, none of Sidney's classroom artifact packet tasks asked students to construct a proof, so it is not possible to assess Sidney's application of the core elements of proof from the samples of student work he provided. In addition, while Sidney considered the tasks in the classroom artifact packets proof tasks, his rubrics did not contain the core elements of proof.

4.3.3.3 Communication of expectations *(RQ3c: In what ways did participants communicate expectations regarding what is required to produce a proof to students?)*

Sidney did not provide much detail on his student work sample cover sheets with respect to how he communicated his expectations to his students:

- "Just explain that they needed to solve the Solve It! Question any way they felt that they should" (Sidney's first classroom artifact packet Cover Sheet, question #2)
- "1st page—work through the task, thinking of ways to solve or answer the questions. Once you have solved it one way, try another" (Sidney's second classroom artifact packet Cover Sheet, question #2)
- "Try to see what they learn to work backwards. The goal was them to see the transition visually to understand what was going on" (Sidney's third classroom artifact packet Cover Sheet, question #2)

When probed for more detail during his post-interview, Sidney said that he did not share his scoring scheme with his students prior to their work on the first and second task, but that he had discussed his expectations in more detail on the third task. This new information was used to adjust the rubric scores for Sidney's Communication of Expectations, as shown in Table 4.28.

Table 4.28 Communication of Sidney's Expectations

Classroom Artifact Packet	Communication Score
First Sample	1
Second Sample	1
Third Sample	3
Average Score	1.6

In summary, it was Sidney's perspective that he was hampered in his efforts to engage his students in reasoning-and-proving by the content of this unit (polynomial operations), the traditional textbook he was assigned, and by the unsupportive shared lesson plans of his colleagues. There might have been other constraints in play, such as weak content knowledge. In addition, Sidney did not engage his students in proof tasks for his polynomial unit, nor did his definition of proof contain all of the core elements. Even if he had found or created a proof task, it is not likely that he would have evaluated his students' products based on all three core elements of proof.

Of the 643 exercises available to Sidney in his textbook for the sections covered in the 15-day data collection period, 33 exercises were reasoning-and-proving exercises (5.1%). Sidney selected 657 exercises during this time period: 227 exercises were done during class, 402 were optional homework exercises, and 28 were assessment exercises. 2.1% of the 657 exercises were reasoning-and-proving exercises, and there is no evidence that Sidney's students spent a significant time on these 14 reasoning-and-proving exercises. Of these 14 exercises, four asked students to make conjectures, two of which were created by Sidney and two of which were modified from the textbook by Sidney. Four of the 14 exercises asked students to correct a mistake in the optional homework.

The tasks in Sidney's classroom artifact packets had the potential to engage some students in novel conjecture-making with respect to developing procedures for factoring and foiling, but the similarity among the students' responses suggested that the students were already familiar with these concepts. Sidney maintained the low-level of cognitive demand between the task potential and implementation for the first task, he showed a decline in the other two tasks which began as high-level.

Sidney's definition of proof was partially based on the core elements of proof. From the work he gave his students, Sidney focused more on the making generalizations rather than on the developing arguments. He communicated some expectations of work to his students. Sidney's rubrics for evaluating his students' work contained elements such as used previous knowledge, sound reasoning, explained why, and completed task individually.

4.4 JONATHAN

Jonathan was a first-year mathematics teacher in a start-up charter school in central Pennsylvania during the study described herein. While the physical location of his school is in a suburban area, the 158 enrolled students come from nine surrounding urban districts and are enrolled in 9^{th} or 10^{th} grade. By the year 2015 the school plans to have ninth through twelfth grades. Jonathan was the only mathematics teacher in the school, but in the 2012-2013 school year the chemistry teacher and a learning support teacher each taught one section of math. Jonathan had complete control over the standards, curriculum, and assessments for mathematics in his school. During his interview, Jonathan expressed that he "spend[s] a lot of hours teaching my fellow math teachers how to teach math, which takes away from my own planning and prep time. But they just have no experience" (personal communication, February 28, 2013).

Jonathan acknowledged the pros and cons of having complete control—but no guidance—over his curriculum (personal communication, February 28, 2013). He chose CCSSM to guide his curriculum and selected Pearson's *Algebra 1: Common Core* (Charles, Hall, Kennedy, Bellman, Bragg, Handlin, Murphy, & Wiggins, 2012) as the textbook for his Algebra 1 course. This is a slightly different version of the same book Sidney used in his Algebra 1 class in Virginia. Because Jonathan's students were new to the school this year, they were not responsible for passing the Keystone Exam (Pennsylvania's state assessment exam), but Jonathan took the content of the Keystone Exam into consideration when he designed his curriculum.

Throughout the year, Jonathan was able to restructure the mathematics courses in his school to accommodate what he described as the shockingly low ability level and lack of mathematical experience of his students. By February, Jonathan reported that his students were still below grade level but were making progress. With respect to engaging his students in reasoning-and-proving, Jonathan reported:

> So it's taken a really long time for me to get them to realize that you have to justify your answers, but we still really struggle, um, with even doing like pattern tasks. They can take two or three days to do just one of them, which I have done just to get them to realize that I'm not just going to give you the answer. So, um, I try to incorporate reasoning-and-proving, but for them right now, reasoning-and-proving is just being able to say why and that's been a huge boundary. (personal communication, February 29, 2013)

In his post-interview on May 29th, 2013, Jonathan was asked how his expectations of engaging his students in reasoning-and-proving had changed in the past year. He reported:

> In the [CORP] class, I loved the idea of giving the students a task and giving the students time to explore in small groups and coming up with something almost entirely on their own. It was my intention to do this at least once a week. With my population this year, that was challenging (no prior work with reasoning-and-proving, no prior independent

exploratory work). I spent a lot of time running around to each group. Throughout the year, I've had to scaffold a lot of the problems. Going forward next year, we will be ready to have a discussion about what it means to reason-and-prove and then take another step in that direction… Ideally, [next year I would] have more tasks where students are reasoning-and-proving on their own. Now students have a shared language and are much more engaged and proud of themselves when working on a task (they don't view this type of activity as a chore, like drill-and-practice work).

The unit Jonathan chose for the current study was linear functions. His 15-day data collection period included the topics rate of change and slope, direct variation, slope-intercept form, point-slope form, standard form, and parallel and perpendicular lines.

4.4.1 Selecting exercises *(RQ1: To what extent did participants select reasoning-and-proving learning opportunities in the form of exercises?)*

4.4.1.1 Available in textbook *(RQ1a: To what extent does the textbook include exercises that have the potential to engage students in reasoning-and-proving?)*

Jonathan's textbook contained 433 exercises in the sections spanned by Jonathan's curriculum during his 15-day data collection period. The textbook offered students exercises to check their understanding of the lesson, exercises for practicing and problem-solving, standardized test preparation exercises, mixed review exercises, and a mid-chapter quiz. Of 433 exercises, 30 had the potential to engage students in reasoning-and-proving, which represents 6.9% of the exercises. Half of the exercises asked the students to investigate a conjecture, and the remainder of the reasoning-and-proving exercises asked students to make a conjecture, correct a mistake, or develop an argument (see Table 4.29).

Table 4.29 Available Types of Reasoning-and-Proving Exercises in Jonathan's Textbook

Type of RP Exercise	Available in Textbook*
Make a Conjecture	7 (1.6%)
Investigate a Conjecture	15 (3.5%)
Evaluate an Argument	0 (0%)
Correct a Mistake	4 (0.9%)
Develop an Argument	5 (1.2%)
Counterexample	0 (0%)
Principles of Proof	0 (0%)
Non-Reasoning-and-Proving	403 (93.1%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

4.4.1.2 Selected by Jonathan *(RQ1b: To what extent did the participant select exercises for instruction that had the potential to engage students in reasoning and proving?)*

Jonathan reported 170 exercises on his Task Log Sheet, 11 of which were judged to be reasoning-and-proving by the coders, representing 6.5% of the exercises. However, if the time spent on these tasks is estimated from Jonathan’s Task Log Sheets, Jonathan’s students spent about 27% of their class time engaged in reasoning-and-proving. The percentage of reasoning-and-proving exercises selected by Jonathan closely matches the percentage of exercises available in Jonathan's textbook (6.9%), but Jonathan offered a more narrow reasoning-and-proving experience than was offered by the textbook (see Figure 4.36). Jonathan offered his students opportunities to make 8 conjectures and investigate 3 conjectures, but did not offer any correct a mistake or develop an argument exercises (see Table 4.30). Notice that Jonathan did not engage his students in any proof exercises during the 15-day data collection period.

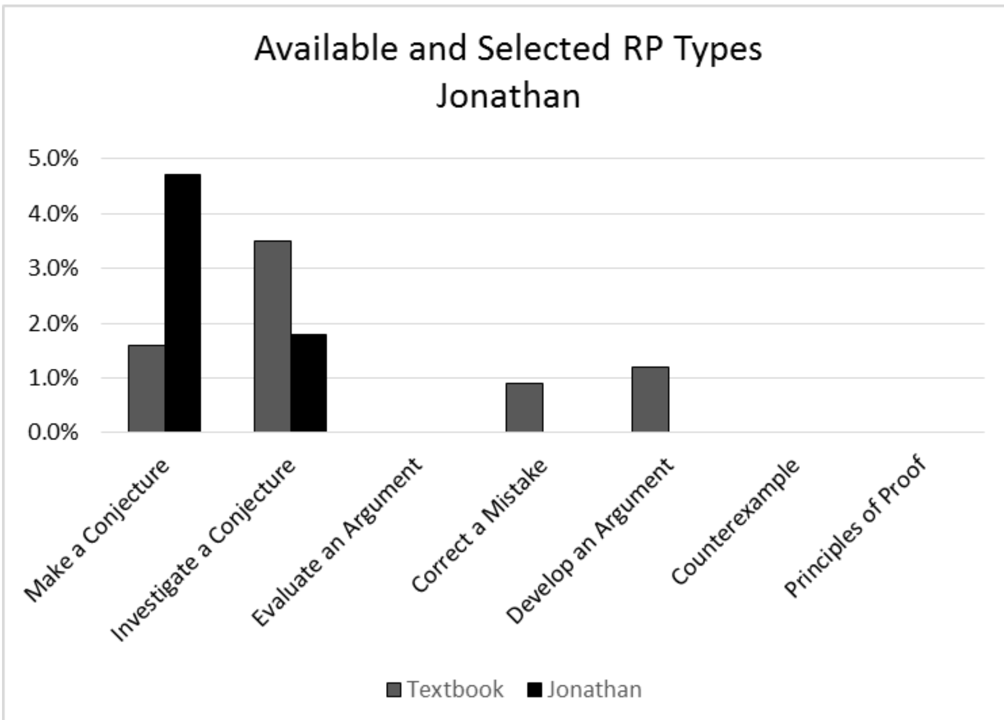


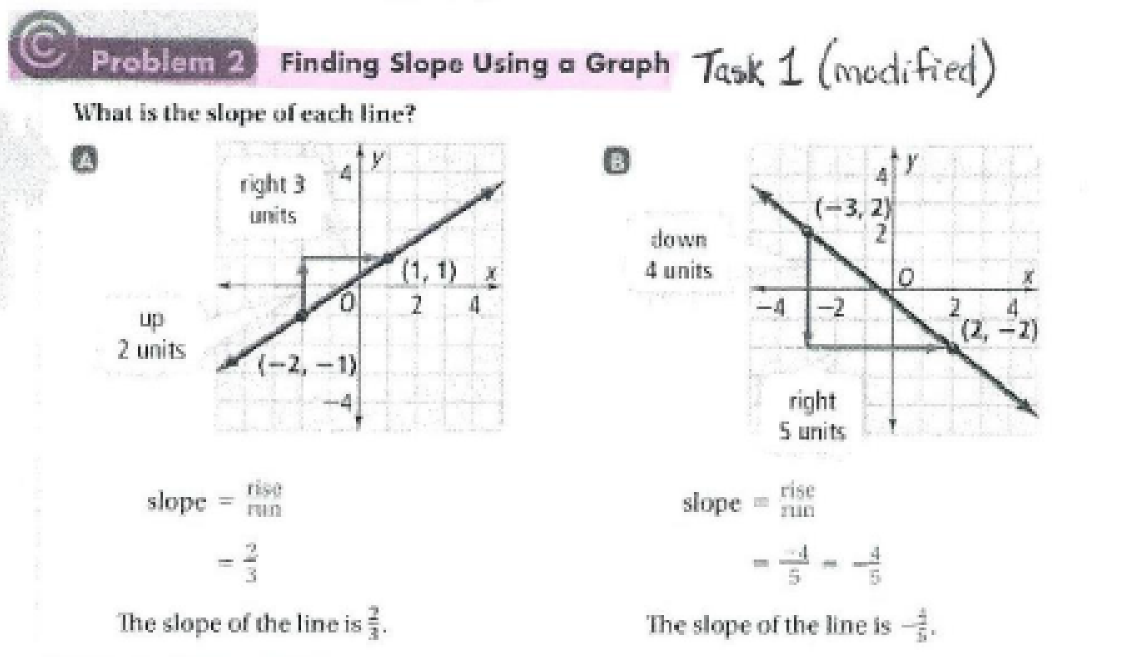
Figure 4.36 Comparison of available and selected types of reasoning-and-proving exercises.

Table 4.30 Selected Types of Reasoning-and-proving exercises by Jonathan

Type of RP Exercise	Selected by Jonathan*
Make a Conjecture	8 (4.7%)
Investigate a Conjecture	3 (1.8%)
Evaluate an Argument	0
Correct a Mistake	0
Develop an Argument	0
Counterexample	0
Principles of Proof	0
Non-Reasoning-and-Proving	159 (93.5%)

**Exercises may be coded as more than one type of reasoning-and-proving, so the percentages may not sum to 100%.*

Of the 11 exercises rated as reasoning-and-proving by the coders, Jonathan labeled 9 as reasoning-and-proving and 2 as not reasoning-and-proving. One exercise he missed labeling as reasoning-and-proving was a modified textbook exercise from Day 1 (Charles et al., 2012, p. 295) involving making a conjecture about linear slopes being constant (see Figure 4.37). The other mislabeled exercise asked students to investigate a conjecture from Day 15 (Charles et al., 2012, p. 330) which is shown in Figure 4.38.



Jonathan's modification:

- Put 2 graphs on ELMO (document camera) (draw on graph paper) and hand out to each table (hand out A 1st, then when finished, hand out B)
- Write question on board and have students answer at tables:

  Pick any point on the line.

  1. From the point you picked, move 1 unit to the right, how many units does the line rise? What happens if you choose a different starting point?
  2. Now move 2 units to the right of your point. How many units does the line rise?
  3. Now move 3 units to the right of your point. How many units does the line rise
  4. What is the slope of this line? How do we find the slope of a line? [Repeat with line B]

Figure 4.37 Mislabeled make a conjecture exercise from Jonathan's lessons (Day 1) (adapted from Charles et al., 2012b, p. 295).

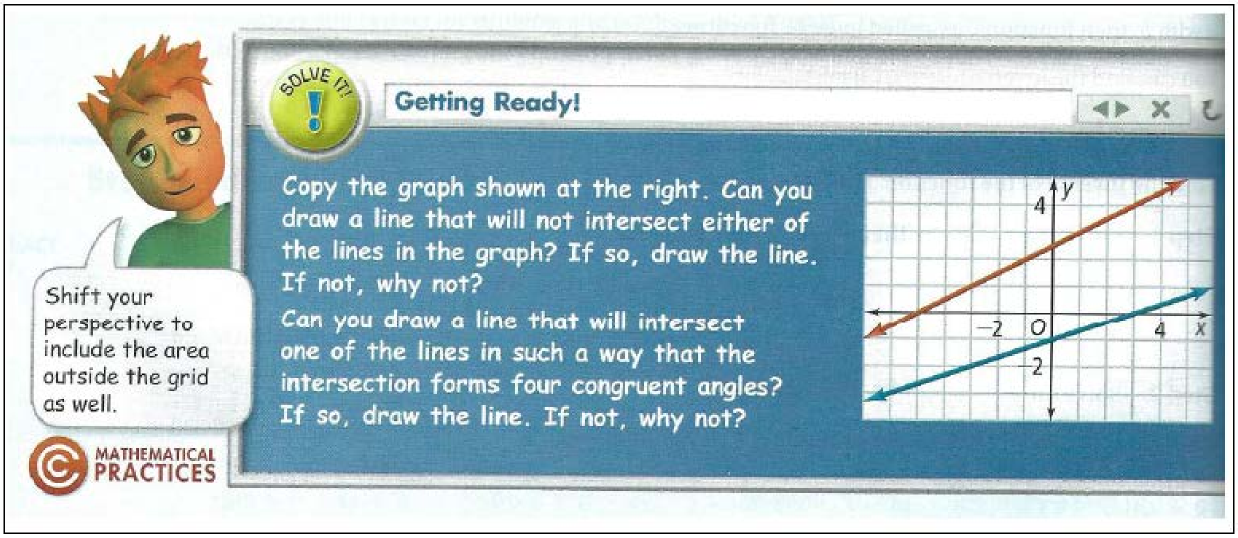


Figure 4.38 Missed investigate a conjecture exercise from Jonathan's lessons (Day 15) (Charles et al., 2012b, p. 330).

There were 159 lesson exercises rated as non-reasoning-and-proving by the coders. Jonathan labeled 33 of these as reasoning-and-proving and 126 as not reasoning-and-proving. Like Uma, however, Jonathan labeled two entire collections of exercises constituting two explorations from Day 5 and Day 15 as "reasoning-and-proving" while the coders separated the exercises and evaluated them individually. On Day 15, for example, Jonathan asked his students to define parallel lines, find the slopes of four pairs of parallel lines, and then make a conjecture regarding when two lines were parallel; the second half of the lesson asked similar questions for perpendicular lines with one additional calculation (products of the slopes). When the 12 exercises in the exploration were rated individually, only the two conjectures counted as reasoning-and-proving. Considering this, Jonathan was consistently able to identify exercises with the potential to engage students in reasoning-and-proving.

4.4.1.3 Exercise modifications *(RQ1c: To what extent did the participant modify exercises to affect the exercises' potential to engage students in reasoning-and-proving?)*

Figure 4.37 shows one example of an exercise modification made by Jonathan to increase its reasoning-and-proving potential. In the course of the 15-day data collection period, Jonathan modified 13 exercises, all from his textbook. Only 2 of the modifications affected the reasoning-and-proving potential of the exercise (see Table 4.31). Both modifications occurred on the first day of data collection; one conjecture involved the fact that linear slopes are constant regardless of where along the line a student determines its value, and the second conjecture was to determine a formula for slope when only given two points (Jonathan did not recognize the first conjecture and reasoning-and-proving but did recognize the second).

Table 4.31 Frequency of Exercise Modifications Made by Jonathan

Original Exercise (any source)	Exercise as assigned by teacher	Code	Frequency
Reasoning-and-Proving Exercise	Exercise assigned, modified to LOWER RP	-1	0
Reasoning-and-Proving Exercise	Exercise assigned, neutral effect of modification	0	0
Reasoning-and-Proving Exercise	Exercise assigned, modified to INCREASE RP	+1	0
Non-Reasoning-and-Proving Exercise	Exercise assigned, modified to INCLUDE RP	+2	2

Jonathan made the following kinds of modifications to the exercises he selected:

- Asked students to explore how to determine slope from any point and explore how to determine slope without counting grid lines (develop a formula for slope when given two points)

- Asked students to determine how to write the equation of a graphed line
- Asked students to determine how to sketch a line from an equation
- Asked students to compare different forms of a line
- Asked students to write the equation of a line in both slope-intercept form and point-slope form and show how to transform one equation into the other

4.4.1.4 Exercise sources (*RQ1d: What were the sources of the tasks that participants selected for instruction?)*

Jonathan largely used his textbook for inspiration and exercises during the 15-day data collection period. Of the exercises taken directly from the textbook, three asked students to investigate a conjecture. The 13 exercises modified from the textbook included the two making conjectures previously described from Day 1. Jonathan used a website once to download a function matching game involving the four representations of linear functions (no reasoning-and-proving present). Finally, Jonathan created over a third of his exercises; six of these offered opportunities for his students to make conjectures. This data is summarized in Table 4.32.

Table 4.32 Sources of Exercises Selected by Jonathan

Source	All Exercises Frequency	Reasoning-and-Proving Exercises Frequency
Taken Directly from Published Textbook/Curriculum	99 (58.2%)	3
Modified from Textbook/Curriculum	13 (7.1%)	2
Used in the CORP Course	0 (0%)	0
Taken or Adapted from Ancillary Resources	1 (0.6%)	0
Created by Teacher	58 (34.1%)	6

4.4.2 Implementation of reasoning-and-proving exercises *(RQ2: To what extent were participants able to maintain the level of cognitive demand of the reasoning-and-proving tasks during implementation?)*

The tasks Jonathan selected for his classroom artifact packets were student explorations of: slope, slope-intercept form, and parallel and perpendicular lines. Figure 4.39 shows the slope exploration from Day 2 of Jonathan's lessons. The goal was for students to connect the value of the slope of a line with the direction of the graph of the line (positive, negative, horizontal, and vertical). Since students were asked to observe a pattern and make a generalized statement, the raters coded this task a 4 for potential of the task. The level of cognitive demand was maintained during implementation, as evidenced by the classroom artifact packet student work samples Jonathan submitted. In these, students made statement such as "Pink: down from left to right; all [slopes] are negative; all negative slopes go down" (sample Jonathan 1-1).

Slope Exploration

Find the slope of the line that goes through each pair of points listed below. Then, graph the line in the color you have chosen for its box.

When you are finished, answer the questions at the bottom of this page. (You can write your answers on the back)

Box Color: ______	Box Color: ______
$(0,0)$ *and* $(3,5)$ Slope	$(-1,4)$ *and* $(2,1)$ Slope:
$(-2,4)$ *and* $(5,18)$ Slope:	$(1,6)$ *and* $(6,4)$ Slope:
$(1,-5)$ *and* $(5,0)$ Slope:	$(-4,2)$ *and* $(2,-3)$ Slope:
Box Color: ______	Box Color: ______
$(-3,2)$ *and* $(2,2)$ Slope:	$(1,7)$ *and* $(1,-3)$ Slope:
$(-5,4)$ *and* $(3,4)$ Slope:	$(0,5)$ *and* $(0,-2)$ Slope:
$(0,7)$ *and* $(-1,7)$ Slope:	$(-4,1)$ *and* $(-4,6)$ Slope:

Analyze your graphs:

1) What do you notice about the graphs of the lines that are the same color?
2) What do you notice about the slopes of the lines that are the same color?
3) Can you make a generalized statement about how the slope of a line is related to its graph?

Figure 4.39 Jonathan's first classroom artifact packet task (Day 2).

In the second student work sample task (Day 5), students were asked to "investigate and explain the behavior of the linear equations in slope-intercept form, $f(x) = mx + b$, as m and b vary" (see Figure 4.40). Students were given series of lines to graph and were asked to record observations and make conjectures about slopes and y-intercepts, then make predictions about the appearance of four new lines. Jonathan was careful to ask students to consider different kinds of values for slope (e.g.,, positive, negative, fraction, whole number). Students recorded conjectures such as, "Positive: line goes up left to right; Negative: line goes down right to left; The bigger the number is the steeper the slope is" (sample Jonathan 2-3).

Slope Intercept Form Exploration and Write-Up

TASK: Investigate and explain the behavior of the linear equations in slope-intercept form,, $f(x) = mx + b$, as m and b vary. To do this, keep one of the values the same and vary the other to see how the graph changes. A good explanation is one that provides a sense of why things work as they do.

What does the graph of the function, $f(x) = mx + b$, look like? Let's investigate this by graphing $f(x)$ with **several values for "m" letting b = 0.**

1) Graph the following lines:
   a. $f(x) = 2x + 0$
   b. $f(x) = x + 0$
   c. $f(x) = \frac{1}{2}x + 0$
   d. $f(x) = -2x + 0$
   e. $f(x) = -\frac{1}{2}x + 0$
   f. $f(x) = -x + 0$
2) Record your observations as m varies in the lines that you graphed above.
3) Make a conjecture about the graph of $f(x) = mx + b$, as "m" changes. How does changing "m" change what the graph looks like? What happens when m is positive? Negative: A fraction? A whole number?
4) Put your conjectures to work. Predict what the graphs of the following look like:

   $f(x) = -3x$ $\quad f(x) = \frac{1}{3}x$ $\quad f(x) = 5x$ $\quad f(x) = x$

   What direction do the lines go? Put them in order from steepest to least steep.

Now let's continue, but **let's keep $m = 1$ and vary b.**

1) Graph the following lines:
   a. $f(x) = x + 2$
   b. $f(x) = x - 2$
   c. $f(x) = x + 3$
   d. $f(x) = x - 3$
   e. $f(x) = x + 1$
   f. $f(x) = x - 1$
2) Record your observations as b varies in the lines that you graphed above:
3) Make a conjecture about the graph of $f(x) = mx + b$, as "b" changes. How does changing "b" change what the graph looks like? What happens when B is positive? Negative?
4) Put your conjectures to work. Predict what the graphs of the following look like:

   $y = 2x + 7$ $\quad y = -5x + 3$ $\quad y = -2x + 2$ $\quad y = \frac{1}{2}x - 1$

Figure 4.40 Jonathan's second classroom artifact packet task (Day 5).

The third task of the student work samples was the parallel and perpendicular line task on Day 15 that was previously described. The task (shown with student work in Figures 4.41 and 4.42) contains prepared examples of perpendicular and parallel lines; students did not have to create their own examples. However, as in Jonathan's other student work samples, students were asked to identify patterns and make conjectures. They were not asked to support their conclusions with mathematical evidence (other than what was provided for them in the tasks) or create a proof. While this level of scaffolding may have been appropriate for the ability level of Jonathan's students, it did make the students' reasoning less explicit. If students had been able to choose their own lines to graph and then see a pattern and group positive slopes together, *et cetera*, with an explanation for mathematical evidence, then the task would have been rated by the coders as a 4 for potential and a 4 for implementation. On his classroom artifact packet cover sheet for the third task, Jonathan wrote:

> I implemented this task as planned, but I was surprised at how many students did not recognize that the slopes of perpendicular lines have a product of -1. Many students quickly found the slopes, but stopped when they had to do multiplication involving a fraction.
>
> I believe this task definitely allowed my students to reach the day's objective (parallel lines have equal slopes and perpendicular lines have slopes that are negative reciprocals/have a product of -1). However, I do not think it took a lot of reasoning on the part of the students. If I implement this again, I may make the worksheet less guided.

Jonathan's reflection indicates an awareness of both his students' current abilities (avoiding multiplication of fractions) and the over-scaffolded nature of the task. It seems that Jonathan had to walk a fine line between accommodating his students' computation needs and developing their reasoning skills. A summary of the potential and implementation of Jonathan's student work samples is shown in Table 4.33.

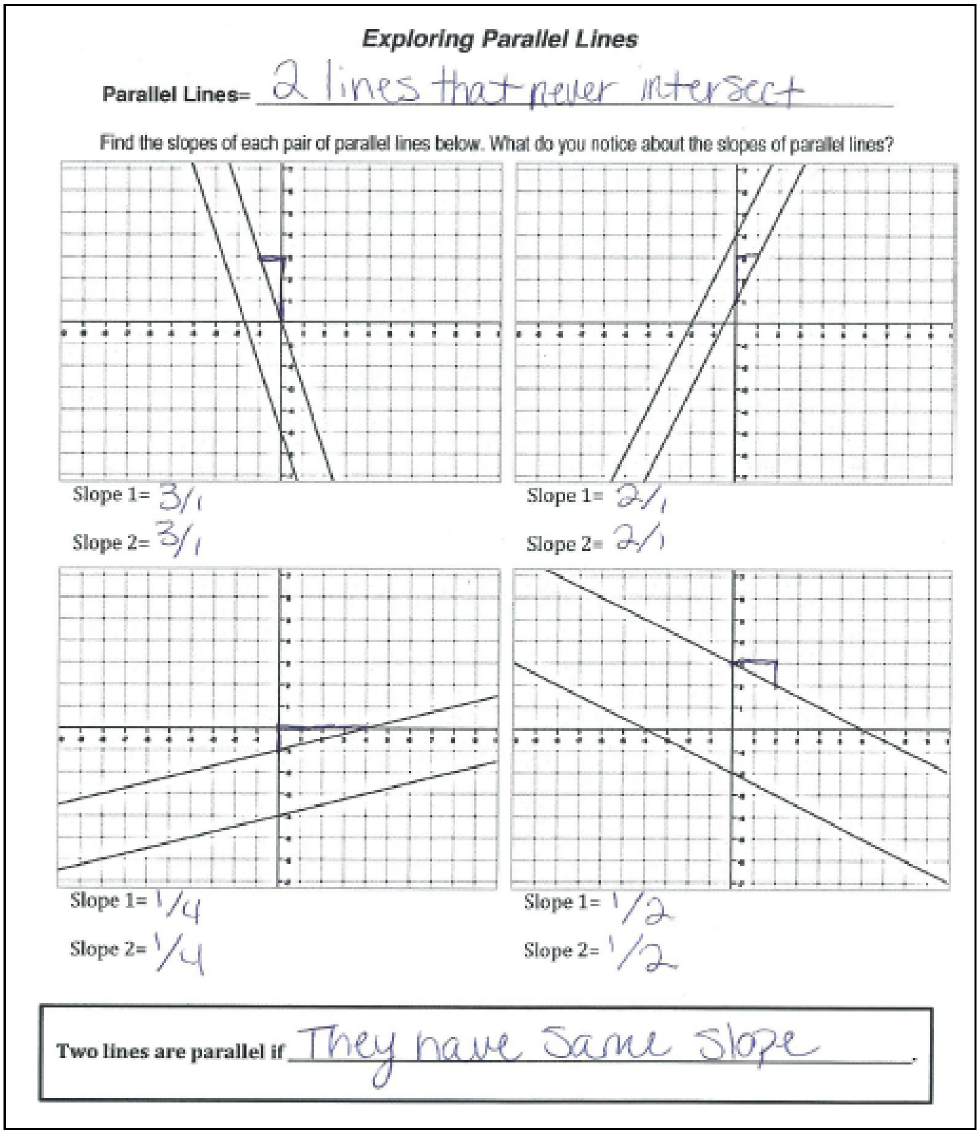
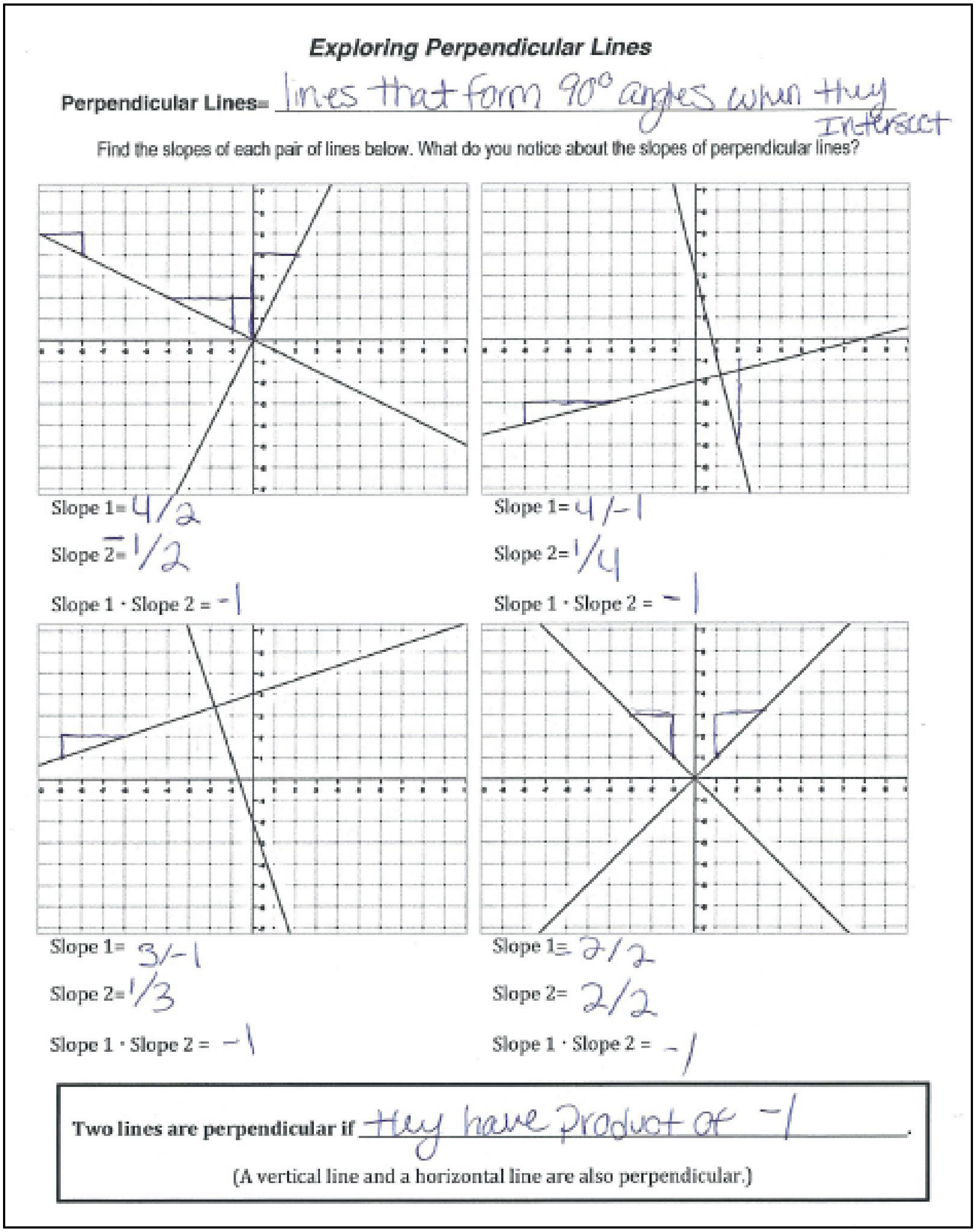


Figure 4.41 Jonathan's third classroom artifact packet task, front page (Day 15).



Exploring Perpendicular Lines

Perpendicular Lines= lines that form 90° angles when they intersect

Find the slopes of each pair of lines below. What do you notice about the slopes of perpendicular lines?

Slope 1= 4/2

Slope 2= 1/2

Slope 1 · Slope 2 = -1

Slope 1= 4/-1

Slope 2= 1/4

Slope 1 · Slope 2 = -1

Slope 1= 3/-1

Slope 2= 1/3

Slope 1 · Slope 2 = -1

Slope 1= 2/2

Slope 2= 2/2

Slope 1 · Slope 2 = -1

Two lines are perpendicular if they have product of -1.

(A vertical line and a horizontal line are also perpendicular.)

Figure 4.42 Jonathan's second classroom artifact packet, back page (Day 15).

Table 4.33 Potential and Implementation of Cognitive Demand of Jonathan's Tasks

Classroom Artifact Packet	Potential of Exercise	Implementation of Exercise	Maintain Cognitive Demand?
First	4	4	Maintain
Second	3	3	Maintain
Third	3	3	Maintain

4.4.3 Evaluation students' reasoning-and-proving products *(RQ3: To what extent were participants able to accurately evaluate their students' reasoning-and-proving products?)*

None of the tasks Jonathan offered to students in his classroom artifact packets were proof tasks, so Jonathan's evaluation of student proof products cannot be assessed. However, Jonathan's post-interview description of proof (personal communication, May 29, 2013) can be compared to the core elements of proof. Jonathan defined proof to a colleague in this way:

> You have this task in front of you, you are trying to come up with an answer. In order to prove something is true, we need to be able to show that it is true in multiple ways and with multiple representations. Be able to tell where all aspects of the formula comes from…connect different representations. Use procedures that have been shown true previously.

Jonathan's description of proof to a student is a little different:

> What does it mean to prove something in an argument (step away from math)? You need evidence that you know your answer is correct. How do you know your evidence is

valid? The evidence must be something you couldn't question. We need to rely on previous definitions and theorems and procedures… if we use those things that we already KNOW are true, then everything we use going forward must also be true. An example is the steps (procedures) in solving an equation; we use the order of operations.

Jonathan captured using statements previously accepted by the community but did not mention a logical flow to the argument. He came close to saying that a mathematical proof must show that a conjecture or claim is (or is not) true for *all* cases but he did not explicitly connect truth and domain under which the statement is true (see Table 4.34).

Table 4.34 Comparison of the Core Elements of Proof with Jonathan's Definition of Proof

Core element of proof	**Jonathan**
The argument must show that the conjecture or claim is (or is not) true for *all* cases.	Partial
The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.	Present
The conclusion that is reached from the set of statements must follow logically from the argument made.	Missing
Additional Criteria: showing a statement is true in multiple ways with multiple representations	Partial

4.4.3.1 Criteria for judgment *(RQ3a: To what extent did participants' criteria for judging the validity of their students' reasoning-and-proving products contain the core elements of proof?)* Jonathan's expectations for his students work were very clear, detailed, and delineated across all three classroom artifact packets (see Table 4.35). In addition, Jonathan wrote on his student work sample cover sheets that the most important aspect of the scoring guide was the quality of the students' conjectures. Jonathan's IQA rubric scores for clarity and detail of expectations are listed in Table 4.36.

Table 4.35 Jonathan's Classroom Artifact Packet Rubrics

Rubric Level	First Rubric	Second Rubric	Third Rubric
Exceeded expectations	All graphs drawn accurately, specifically describes the relationship between the lines in each box by describing the direction, correctly describes the type of slope that each box has in common (positive, negative, zero, or undefined), makes generalized, detailed statements about how the direction of a linear graph is related to its slope.	All questions answered completely and accurately, makes observations about how both *m* and *b* effect the graph of the line and uses these observations to make generalized conjectures about the variables, conjectures state how *m* effects the slope and *b* effects the y-intercept of the graph, and both "predict what the graphs will look like" questions are answered with only 0-1 errors.	Parallel and perpendicular lines are defined correctly, slopes of all 8 lines on both sides of the sheet are calculated correctly, conjectures are made stating that parallel lines have the same slope and perpendicular lines' slopes have a product of -1, conjectures are accurate, concise, and use appropriate vocabulary (e.g., uses "product of -1" as opposed to "multiplied together equals -1").
Met expectations	All graphs drawn accurately, generally describes the direction of the lines in each color, correctly describes the type of slope that each box has in common (positive, negative, zero, or undefined), makes generalized statements about how the direction of a linear graph is related to its slope, but the statements are lacking detail	Most questions answered completely and accurately with the exception of 1 or 2, makes a few, brief observations about how both *m* and *b* effect the graph of the line and uses these observations to make a generalized conjecture about at least one of the variables, conjectures state how *m* effects the slope and *b* effects the y-intercept of the graph, but proper vocabulary is not always used (e.g., "where it goes through the y-axis" instead of "the y-intercept of the graph", and both	Parallel and perpendicular lines are defined correctly, slopes of all 8 lines on both sides of the sheet are calculated correctly, conjectures are made stating that parallel lines have the same slope and perpendicular lines' slopes have a product of -1, conjectures are accurate, but may be a bid wordy and may not use appropriate vocabulary.

		"predict what the graphs will look like" questions are answered with only 2-4 errors.	
Failed expectations	Most graphs are drawn accurately, does not accurately describe the relationship between the lines from each box, does not make accurate statements about the type of slope each box has in common, does not make generalized statements about how the direction of a linear graph is related to its slope	Many questions are not answered or are partially answered, makes a few observations about how *m* and/or *b* effect the graph of the line, but observations are vague and/or inaccurate, conjectures are not provided or are incorrect, and students are unable to complete the "predict what the graphs will look like" questions, or they complete them incorrectly.	Parallel and perpendicular lines are defined, but definitions are not entirely correct, slopes are not all found or are not all correct, there is no evidence of conjectures made, or the conjectures are incorrect.

Table 4.36 Quality of Jonathan's Rubrics Used to Judge Student Work

Classroom Artifact Packet Rubric	Clarity and Detail of Expectations	Comment
First	4	*Detailed, clear, delineated, and emphasis on quality of conjecture*
Second	4	*Detailed, clear, delineated, and emphasis on quality of conjecture*
Third	4	*Detailed, clear, delineated, and emphasis on quality of conjecture*
Average	*4*	

4.4.3.2 Communication of expectations *(RQ3c: In what ways did participants communicate expectations regarding what is required to produce a proof to students?)*

Jonathan described his communication of expectations both in terms of students' behavior and mathematical products. In the first classroom artifact packet, he wrote, "Students were instructed to work in pairs. In regards to the 'analyze your graphs' questions, I prompted students to be as descriptive as possible in their descriptions." For the second student work sample, Jonathan wrote,

> Before beginning the task, I walked them through the process of graphing on this calculator. I also reinforced 'A good explanation is one that provides a sense of why things work as they do,' which is included in the task directions on the worksheet. (Jonathan's second classroom artifact packet Cover Sheet)

Jonathan was less clear about his communication of expectations for the third task on the classroom artifact packet cover sheet, stating only that he helped the class come to a collaborated definition of parallel and perpendicular lines. Since Jonathan's students completed open-ended tasks of this type about once a week, it was assumed that the students were familiar with

Jonathan's expectations of high-quality work and had seen many examples in previous class sessions. Therefore, Jonathan received a communication score of 4 for discussing and modeling his expectations to his students (see Table 4.37).

Table 4.37 Communication of Jonathan's Expectations

Classroom Artifact Packet	Communication Score
First	4
Second	4
Third	4
Average Score	4

In summary, Jonathan offered his students a similar proportion of reasoning-and-proving opportunities as did his traditional Algebra 1 textbook (6.5% compared to 6.9%). He showed a good recognition of reasoning-and-proving exercises, and he created 58 of the 159 exercises he offered to his students. These created exercises included 6 of the 11 reasoning-and-proving exercises with which he engaged his students. Jonathan's reasoning-and-proving exercises were limited to making and investigating conjectures.

Jonathan maintained the cognitive demand of the three tasks he submitted with his classroom artifact packets and showed an awareness of the level of scaffolding included in the problems that prevented his students from truly engaging in cognitively-demanding reasoning-and-proving work. However, Jonathan's scoring rubrics for his students' work were focused on the core mathematical content of the task and showed a high level of detail.

4.5 TRENDS—LOOKING ACROSS THE FOUR CASES

4.5.1 Selection of Exercises *(RQ1: To what extent did participants select reasoning-and-proving learning opportunities in the form of exercises?)*

The data in Table 4.38 shows a sharp contrast between the percent of reasoning-and-proving exercises available in the reform curriculum (75.5%) compared to any of the traditional curricula (average 5.7%) and a contrast among the participants with respect to the number of and time spent on selected reasoning-and-proving exercises. There was not a comparable difference between the traditional geometry and the traditional algebra 1 curricula (5.2% average compared to 6%). Some of these differences could be attributed to the content which was selected for the 15-day data collection period. Karen chose a unit on proof and congruence, while Uma chose a unit on area and surface area. Sidney chose a unit on polynomial operations, while Jonathan chose a unit on linear functions. Obviously, a unit devoted to proof and congruence should contain a high percent of reasoning-and-proving exercises.

Table 4.38 Percentage of Available and Selected Reasoning-and-proving Exercises

Participant	Course / Curriculum Type	Percentage of RP Exercises Available in Textbook	Percentage of RP Exercises Selected by Participant	Estimated Time Spent on RP Exercises during Lessons
Karen	Geometry/reform	75.5%	76.4%	76%
Uma	Geometry/traditional	5.2%	3.4%	16%
Sidney	Algebra 1/traditional	5.1%	2.1%	4%
Jonathan	Algebra 1/traditional	6.9%	6.5%	27%

The types of reasoning-and-proving exercises available in the textbooks used by participants during their 15-day data collection period is represented in Figure 4.43. Looking across all of the available reasoning-and-proving exercises, most asked students to investigate conjectures or develop arguments (averages of 43.9% and 38.6% respectively). Making conjectures were represented in 21.2% of the exercises, but very few exercises asked students to evaluate arguments (1.8%), find counterexamples (3.9%), or think about the principles of proof (5.5%). If Karen's curriculum is removed, the type available most frequently is still investigate a conjecture (37.5%), followed by make a conjecture (34.6%), develop an argument (16.3%), correct a mistake (12.5%), evaluate an argument and find a counterexample (3.8% and 2.9% respectively), and principles of proof (0%).

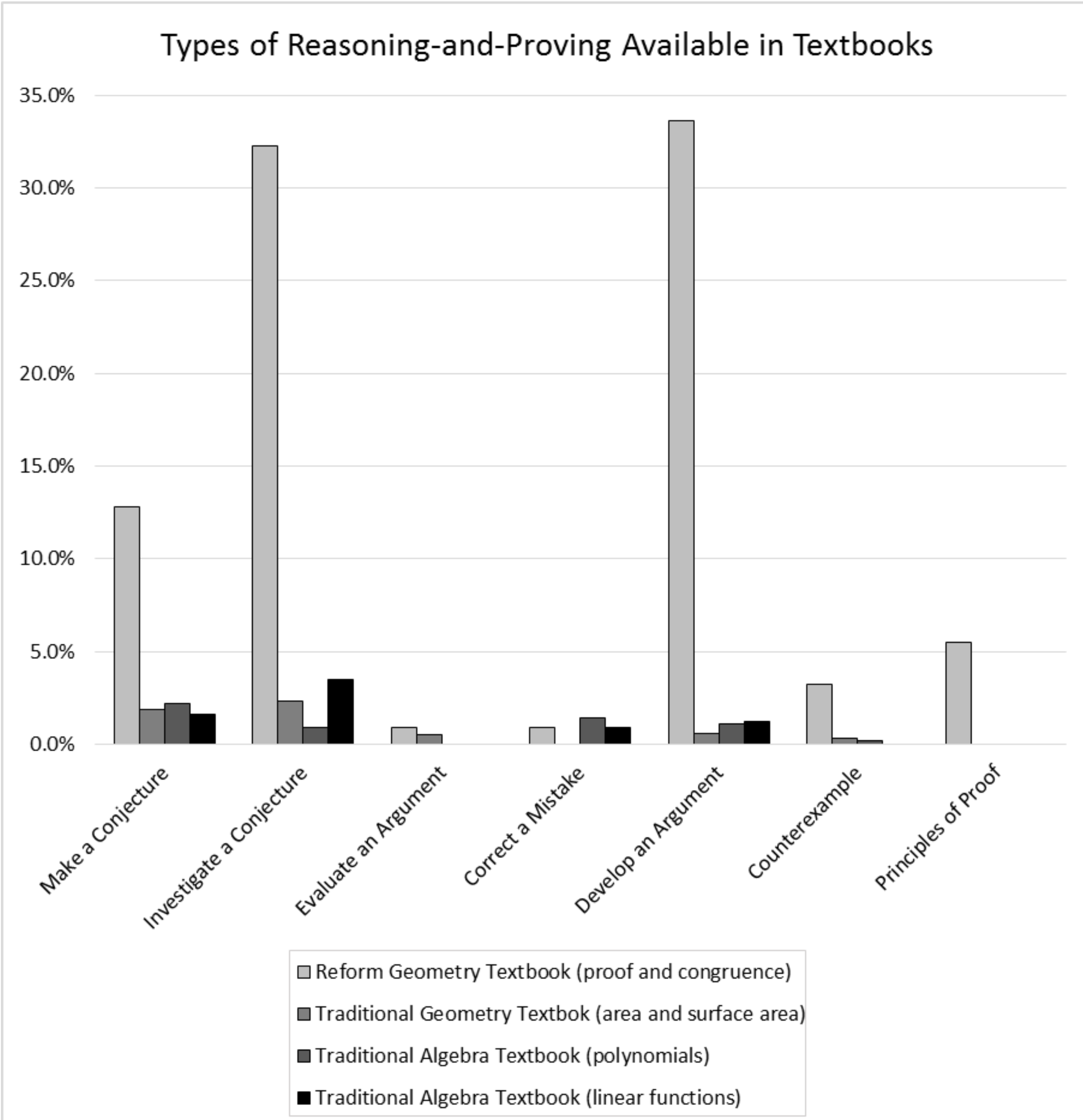


Figure 4.43 Types of reasoning-and-proving exercises in textbooks used by participants during their 15-day data collection period.

The types of reasoning-and-proving exercises selected by participants show a preference towards making and investigating conjectures and developing arguments (see Table 4.39). With the exception of Karen, no one selected any Evaluate an Argument or Principles of Proof exercises. Correct a Mistake and find a Counterexample types of exercises were poorly represented as well. This trend is similar to the types of reasoning-and-proving available in the participants' textbooks.

Table 4.39 Types of Reasoning-and-proving Exercises Selected by Participants During Their Data Collection Period

Reasoning-and-Proving Type	Karen	Uma	Sidney	Jonathan
Make a Conjecture	7.2%	2.8%	0.8%	4.7%
Investigate a Conjecture	49.6%	0.7%	0.6%	1.8%
Evaluate an Argument	0.4%	0.0%	0.0%	0.0%
Correct a Mistake	0.8%	0.0%	0.6%	0.0%
Develop an Argument	48.8%	1.7%	0.3%	0.0%
Counterexample	6.3%	0.3%	0.0%	0.0%
Principles of Proof	7.0%	0.0%	0.0%	0.0%

Table 4.40 shows the core elements of proof in the participants' definition and the total number of exercises in the participants' lessons compared to the number of these exercises that were reasoning-and-proving, the number of exercises that were correctly identified as reasoning-and-proving, and the number of exercises that were mislabeled. Karen correctly identified 94.9% of her reasoning-and-proving exercises, while Uma identified 90% of hers, Sidney 71.4% of his, and Jonathan 81.8 % of his exercises. The participants' ability to identify reasoning-and-

proving tasks does not appear to be correlated with how many core elements of proof were in their definition of proof.

Table 4.40 Frequency and Identification of Reasoning-and-proving Exercises Selected by Participants

Participant	Number of Core Elements of Proof (maximum of three)	Number of Exercises in Lessons	Number of RP Exercises in Lessons	Number of RP Lesson Exercises Labeled by the Participant in Lessons	Number of Lesson Exercises Labeled by the Participant as RP that Were Not RP
Karen	2.5	335	256	243	77
Uma	2	290	10	9	32
Sidney	1.5	657	14	10	19
Jonathan	1.5	170	11	9	33

The number of exercises that participants mislabeled as reasoning-and-proving can largely be explained by labeling collections of exercises rather than individual exercises. In Uma and Jonathan's cases, the mislabeled exercises were parts of larger tasks that would be labeled collectively as reasoning-and-proving tasks. In other words, some of the subparts of the tasks, such as generating data in order to find a pattern, would not individually be labeled as reasoning-and-proving exercises but they were in the service of helping students see patterns and make conjectures. Otherwise, Karen and Sidney seemed to mislabel exercises that involved reasoning about definitions or procedures but not the kind of reasoning that leads to proof.

The small number of reasoning-and-proving tasks available and selected by participants makes it difficult to identify trends in types of exercises that are difficult to recognize by participants (see Table 4.41). If a type of reasoning-and-proving was not selected by the participant, the participant did not have the opportunity to label it reasoning-and-proving or not, so "not applicable" was entered in the data table. Correct a Mistake, Develop an Argument, and Counterexample types of exercises appear to be easily identifiable as reasoning-and-proving because none of the participants missed labeling them. Making a Conjecture and Identifying a Conjecture appear to be more difficult to recognize, especially for Karen (although she correctly identified 20 other Make a Conjecture exercises) and Sidney (missed all four Investigate a Conjecture exercises he selected). Principles of Proof only applied to Karen; this type must be unexpected or difficult to recognize because Karen missed identifying half of these types of reasoning-and-proving exercises.

Table 4.41 Types of Reasoning-and-Proving Exercises that Participants Missed Identifying

Reasoning-and-Proving Type Missed	Karen	Uma	Sidney	Jonathan
Make a Conjecture	4	0	0	1
Investigate a Conjecture	0	1	4	1
Evaluate an Argument	0	n/a	n/a	n/a
Correct a Mistake	0	n/a	0	n/a
Develop an Argument	0	0	0	n/a
Counterexample	0	0	n/a	n/a
Principles of Proof	9	n/a	n/a	n/a

Table 4.42 shows the modifications participants made on reasoning-and-proving exercises or on exercises that were modified to include reasoning-and-proving. Note that participants made modifications to non-reasoning-and-proving exercises in their lessons, but those modifications are not represented here unless the exercise became a reasoning-and-proving exercise with the modification. No participant lowered the reasoning-and-proving potential of an exercise through modification. A total of 29 modifications were made to the wording in reasoning-and-proving exercises which did not affect the exercises' potential to engage students, and 7 modifications resulted in an increase in an exercises reasoning-and-proving potential. The burden of increasing the potential of exercises fell to Uma, Sidney, and Jonathan, the three participants using traditional textbooks. In addition, Sidney and Jonathan—Algebra 1 teachers—modified 5 exercises which were not originally reasoning-and-proving at all. These represent 26.3% of the reasoning-and-proving exercises in which they engaged their students.

Table 4.42 Frequency and Effect of Modifications Made to Exercises by Participants

Original Exercise (any source)	**Exercise as assigned by participant**	**Frequency of Modifications**			
		Karen	**Uma**	**Sidney**	**Jonathan**
Reasoning-and-Proving Exercise	Exercise assigned, modified to LOWER RP	0	0	0	0
Reasoning-and-Proving Exercise	Exercise assigned, neutral effect of modification	27	1	0	0
Reasoning-and-Proving Exercise	Exercise assigned, modified to INCREASE RP	0	0	0	0
Non-Reasoning-and-Proving Exercise	Exercise assigned, modified to INCLUDE RP	0	2	3	2

The data in Table 4.43 shows that Uma and Jonathan created half of the reasoning-and-proving exercises with which they engaged their students. Uma created 5 of the 10 reasoning-and-proving exercises she selected and Jonathan created 6 of the 11 reasoning-and-proving exercises he selected. Sidney created 2 exercises; the other 12 reasoning-and-proving exercises he selected came directly from the textbook or from a modified version of a textbook exercise. Karen, who had 330 reasoning-and-proving exercises available in her textbook, still created 10 reasoning-and-proving exercises on her own.

Table 4.43 Sources of Reasoning-and-proving Exercises Selected by Participants

Source	**Frequency of Reasoning-and-Proving Exercises**			
	Karen	**Uma**	**Sidney**	**Jonathan**
Taken Directly from Published Textbook/Curriculum	188	2	10	3
Modified from Textbook/Curriculum	27	3	2	2
Used in the CORP Course	0	0	0	0
Taken or Adapted from Ancillary Resources	32	0	0	0
Created by Teacher	9	5	2	6

4.5.2 Implementation of reasoning-and-proving exercises *(RQ2: To what extent were participants able to maintain the level of cognitive demand of reasoning-and-proving tasks during implementation?)*

Table 4.44 summarizes the average potential and implementation scores for the classroom artifact packets each participant submitted. Karen, Uma, and Jonathan were able to successfully select and implement cognitively-challenging reasoning-and-proving tasks. Karen's first classroom artifact packet contained the first proof activity her students were asked to complete; it is no surprise that some statements (e.g., givens) were missing in some students' proof attempts and these minor omissions did not prevent students' attempts from exceeding Karen's expectations. Such omissions, however, prevented those arguments from being assessed as proofs (her implementation was thus rated a 3 rather than a 4, which is still considered maintaining the level of cognitive demand of the task). Jonathan recognized that his tasks were overly scaffolded with respect to the level of cognitive demand and plans to remove some scaffolding next year to provide his students with more opportunity to grapple with reasoning-and-proving. Uma's tasks were the most different from each other; her classroom artifact packet tasks spanned a statement to disprove (requiring only a counterexample), an irregular area task (which left little ambiguity about what needed to be done), and a maximum area for a fixed perimeter exercise for which students were supposed to justify why a circular animal pen was the correct answer.

Table 4.44 Average Potential and Implementation Scores for Cognitive Demand of Tasks in Participants' Classroom Artifact Packets

Participant	Average Exercise Potential Score	Average Exercise Implementation Score	Average Maintenance
Karen	4	3.6	Maintained
Uma	3.3	3	Maintained
Sidney	2.7	2	Declined
Jonathan	3.3	3.3	Maintained

4.5.3 Evaluation of students' reasoning-and-proving products *(RQ3: To what extent were participants able to accurately evaluate their students' reasoning-and-proving products?)*

All four participants described proof as using statements, definitions, and theorems that had been previously validated or accepted by the students in the classroom. The participants also described proof as a way to validate a claim, but in most cases, omitted any reference to the domain under which this truth holds. Only one participant (Karen) mentioned anything about the logical flow of the argument, perhaps because her unit focused on teaching students how to write proofs. Table 4.45 summarizes these results.

Table 4.45 Comparison of the Participants' Definition of Proof with the Core Elements of Proof

Core Elements of Proof	**Karen**	**Uma**	**Sidney**	**Jonathan**
The argument must show that the conjecture or claim is (or is not) true for *all* cases.	Partial	Present	Partial	Partial
The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.	Present	Present	Present	Present
The conclusion that is reached from the set of statements must follow logically from the argument made.	Present	Missing	Missing	Missing
Additional Criteria	None	None	Partial	Partial

Participants were consistent in their clarity and detail of expectations (see Table 4.46). Karen and Jonathan focused on delineating mathematical understanding in their levels. Specifically, Karen used the core elements of proof to distinguish performance for her students, and Jonathan placed emphasis on the quality of his students' conjectures. Uma and Sidney broadly based high, medium, and low performance partially on behavior.

Table 4.46 IQA Rubric: Clarity and Detail of Expectations Scores

Student Work Sample	**Karen**	**Uma**	**Sidney**	**Jonathan**
First	4	3	3	4
Second	4	3	3	4
Third	4	2	3	4
Average	*4*	*2.7*	*3*	*4*

Participants were largely consistent with how they communicated their expectations to their students as well (see Table 4.47). Karen's students regularly discussed and attempted proof tasks, and Karen orally and in writing reminded students of her specific expectations with respect to their behavior and mathematical work. Uma's expectations were clear but her students did not practice these kinds of tasks very often, so they did not have many models of high-quality work. Sidney said that he did not share his expectations with his students until the third task, before which he engaged his students in conversation about expectations. Jonathan claimed he did not model high-level work for each task prior to students engaging in the tasks, but he also wrote that his students work on exploratory tasks once a week, which implies that students had indeed seen models.

Table 4.47 IQA Rubric: Communication of Expectations Scores

Classroom Artifact Packet	Karen	Uma	Sidney	Jonathan
First	4	3	1	4
Second	4	3	1	4
Third	4	3	3	4
Average	*4*	*3*	*1.7*	*4*

This chapter addressed the three main research questions for this study (selection, implementation, and evaluation of student work products of reasoning-and-proving exercises by participants) using the analyses presented in Chapter Three. The results and analysis were presented in the form of a narrative case for each participant, followed by a trend analysis. Chapter Five explores the possible reasons why some teachers were more successful than others in selecting, implementing, and evaluating the student work products of reasoning-and-proving exercises and tasks by participants.

5.0 DISCUSSION

In this chapter, a discussion of this investigation and how the results inform the extent to which trained teachers selected, implemented, and evaluated tasks in order to offer their students reasoning-and-proving learning opportunities in their classes is presented. The chapter begins by describing the importance of this study. This is followed by a discussion of each of the main research questions, then a discussion of the possible reasons for the results. The chapter closes with concluding remarks and suggestions for future research.

5.1 IMPORTANCE OF THE STUDY

This study examined the extent to which four early career teachers selected, implemented, and evaluated reasoning-and-proving tasks and products in their classrooms after having received training in the form of a preservice university course on reasoning-and-proving. The CORP course content helped participants learn about the need for proof, the core elements of proof, ways to support students, and the connections between tools in an effort to address common problems associated with the implementation of reasoning-and-proving (concept of proof, insufficient resources, and pedagogical challenges). The intended outcomes of the course included an understanding of what constitutes reasoning-and-proving, an understanding of how secondary students benefit from engaging in reasoning-and-proving, and an understanding of how they can support the development of students' capacities to reason-and-prove. Initial evidence from pilots conducted at three large research universities suggest that preservice teachers engaged with the materials, believe that what they learned in the course had the potential to impact their teaching practice, and that the cases helped preservice teachers think about instructional issues related to reasoning-and-proving. In addition, the results suggest that the tasks-tools-talk structure was a useful framework for supporting preservice teachers' analysis and discussion of the cases (Smith et al., 2012). The importance of the study reported herein is that it follows teachers who participated in the course into their classrooms, where external pressures such as district curricula and live students effect the teachers' ability to enact reasoning-and-proving. This study provides an opportunity to see how teachers operationalized their content knowledge and pedagogical content knowledge of reasoning-and-proving.

5.2 SELECTION, IMPLEMENTATION, AND EVALUATION OF REASONING-AND-PROVING TASKS AND PRODUCTS

5.2.1 Selection of reasoning-and-proving tasks

The results of this study suggest that the participants were able to select and create exercises that had the potential to engage students in reasoning-and-proving. In the CORP course, participants solved many reasoning-and-proving tasks that offered a variety of learning opportunities. Participants also learned to modify insufficient tasks to improve the tasks' potential to engage students in reasoning-and-proving. G. J. Stylianides' (2010) analytical framework of reasoning-and-proving (see Figure 2.1) helped participants to situate tasks along a continuum of scaffolded activities that help students develop the notion of and ability to create proofs. This body of knowledge that the participants learned in the CORP university course is particularly important due to the lack of reasoning-and-proving found in most textbooks (Lithner, 2004; Stylianides, G. J., 2009; Thompson, Senk, & Johnson, 2012).

Uma, Sidney, and Jonathan's textbooks contained few exercises designed to engage students in reasoning-and-proving. The small number of reasoning-and-proving exercises available in their traditional textbooks mirrored Thompson, Senk, and Johnson's (2012) study. Not only were the number of exercises small, but the types of reasoning-and-proving exercises were not evenly represented. Very few exercises asked students to evaluate an argument, correct a mistake, find a counterexample, or think about principles of proof. The only textbook which had a significant number of reasoning-and-proving exercises and spanned all types was Karen's CME *Geometry: Unit 2: Congruence and Proof* book. However, based on Karen's description of the reasoning-and-proving opportunities offered to her students beyond this unit, it appears that CME treated proof as a stand-along topic rather than a practice that spanned the entire course.

Despite the limitations of their textbooks, Karen, Uma, and Jonathan spent a significant amount of time engaging students in reasoning-and-proving (76%, 16%, and 27%, respectively). Sidney spent the least amount of time on reasoning-and-proving (4%) and offered the fewest number of opportunities (2.1%). All of the participants were able to recognize and create reasoning-and-proving exercises, with Karen and Sidney taking a broad view of reasoning-and-proving and over-labeling some exercises. This finding is significant because research has shown that secondary students spend very little instructional time on reasoning-and-proving (Porter, 1993). What is still unclear is the extent to which the time spent on reasoning-and-proving is typical for the participants or atypical because they were asked to choose a unit in which they would be asking their students to engage in reasoning-and-proving.

Knuth (2002c) described five roles of proof: to verify that a statement is true, to explain why a statement is true, to communicate mathematical knowledge, to discover or create new mathematics, or to systematize statements into an axiomatic system. The four teachers in the study described herein were asked to submit three classroom artifact packets of reasoning-and-proving tasks, one of which had to be a proof task. The results varied; Karen selected three two-column geometry proofs, Uma selected one disproof (find a counterexample to prove a statement false), one procedure-type activity, and one geometry proof (not two-column); Sidney selected three activities that asked students to make conjectures about polynomial operations; and Jonathan asked students to explore and write conjectures about slopes and intercepts of lines.

If all of the activities students completed during the 15-day data collection period are examined, it seems that Karen selected proof tasks to help communicate knowledge and help her students verify truth. Her textbook states, “mathematical proof is a method that relies on certain assumptions, precise definitions, and logical deductions to prove new facts” (CME *Geometry: Unit 2*, p. 71). Karen’s students explored congruence, angle relationships in sets of parallel lines cut by a transversal, generalizing a conjecture which was based on empirical evidence, writing proofs, and quadrilateral properties. While the CME authors approached developing an axiomatic system by the order of the topics in this unit, it is unclear whether that was a focus of Karen’s instruction. Most of the proofs that Karen’s students completed were in a two-column format, historically used to make grading easier and not necessarily to foster understanding

(Herbst, 2002). Students' two-column proofs are sometimes evidence of learning trivial and uninteresting propositions (Herbst, 2002). Herbst suggested that this type of proof-making has limited ability to construct knowledge for the student. Since Karen's students did not create formal proofs after the completion of this unit (personal communication, February 20, 2013), it is possible that Karen's students did not see proof as a vehicle for sense-making. The types of proof—triangle congruence—in Karen's selected unit were somewhat limited in their ability to explicitly reveal students' reasoning. Geometry students frequently cut their teeth on triangle congruence proofs because they are small and manageable. A simple proof can provide students with three givens, and the students just have to list these givens and select the type of triangle congruence that fits the givens (e.g.,, side-angle-side, angle-angle-side, side-side-side, or angle-side-angle). Consequently, many of these proofs when written in two-column format are four lines long and everyone's proof looks the same. Therefore, a teacher runs the risk of misinterpreting student proofs as indicating understanding as opposed to mimicking a formulaic procedure. That being said, a student still has to choose which type of triangle congruence fits the situation and has to know to list the givens first in setting up the logical argument. So despite the fact that Karen's students were able to successfully complete three proofs, it is difficult to assess what Karen knows and understands about the value of proof and her students' abilities to construct proof from the three classroom artifact packets she submitted.

Uma admitted that she did not spend much time on reasoning-and-proving in her geometry class because she is constrained by her state curriculum and the other teachers in her building. Based on her classroom activities, it appears that Uma used proof to verify truth, explain why something was true, and communicate knowledge. Her students developed formulas for the areas of may two-dimensional shapes and the surface areas of many three-dimensional shapes; these formulas were always based on previously learned formulas (e.g., the surface area of a prism was based on the areas of triangles and polygons), but with the amount of scaffolding in her lessons, it is questionable how much reasoning and critiquing her students learned to do on their own. However, Uma's two proof tasks do reveal that she can create tasks that reveal understanding of congruence and counterexamples and require students to write persuading arguments.

Jonathan also had students who historically struggled in mathematics like Uma. He did not use proof tasks but he did consciously helped his students foster reasoning skills in explaining why something was true and to communicate new knowledge. Jonathan was the only participant who articulated that he was laying the groundwork for his students to develop reasoning-and-proving skills that would be used in the future. In Jonathan's case—as the only mathematics teacher in his school—he was guaranteed to teach the same students the following year and could continue his work. As such, Jonathan had investment opportunities with his students that the other participants did not. Jonathan gave his students opportunities to work in groups, explore open-ended tasks, develop a shared vocabulary, view mathematics as something beyond a set of procedures, and make conjectures. By encouraging and holding students accountable for using specifics in their descriptions and conjectures, Jonathan was refining his students' attention to detail which should serve them well as they develop their ability to reason-and-prove.

Sidney is an interesting case. It appears that he engaged his students in the fewest number of reasoning-and-proving exercises and spent the least amount of time on them. The coding system used to analyze Sidney's classroom activities was from Thompson, Senk, and Johnson (2012), who looked separately at the narrative and practice sections of textbooks. Since observations of the four participants were not conducted, the coding system for exercises (e.g., make a conjecture, evaluate an argument) was applied to the in-class activities and the homework activities of the participants in the 15-day data collection period. Through this lens, Sidney's students had little opportunity to engage in reasoning-and-proving. Through the *narrative* lens of Thompson, Senk, and Johnson—which examines the non-practice descriptions of content in textbooks—the picture of Sidney's class changes. Thompson, Senk, and Johnson "believe that the narrative [portion of textbooks] provides opportunities for teachers to introduce reasoning and proof to students" (p. 255) and chose to code how textbooks justified properties according to (p. 261):

- The property is justified with a proof
- The property is justified using a deductive argument based on a specific case or cases
- A justification of the property in the exercises for which a justification of some type is required
- There is no justification provided and no explicit mention is made of leaving the justification to the student

The researchers listed properties which warranted justification for exponents, logarithms, and polynomials in Algebra 1, Algebra 2, and Precalculus textbooks. Sidney helped his students justify the following properties listed by Thompson, Senk, and Johnson: multiplying binomials, squaring a binomial, multiplying to get the difference of squares, factoring quadratic trinomials, factoring perfect square trinomials and the difference of two squares, and finding binomial factors. Thus, how Sidney engaged his students in justifying polynomial properties merits consideration.

10.4 Multiply two binomials

The rectangle has a width of $(x+2)$ and a height of $(2x+1)$.

1.) Draw a square with side length x.
    a. Label the sides.
    b. What would the area of the square be?

2.) Draw the model of the rectangle with the width of $(x+2)$ and a height of $(2x+1)$
    a. Label each side length of the rectangle.
    b. What is the area of each part of the rectangle?
    c. Label each area.

3.) Find an expression for the total area of the rectangle by adding together the areas of the parts.
4.) Use an area model to multiply $(x+3)$ and $(2x+4)$.

Example 1: Find the product of $(x+8)(x-7)$ **Punnet Square method:**

"Algebra Tiles"

Example 2: Find the product of $(2x+3)(5x+1)$.

"Distribution Method:

Example 3: Find the product of $(x-4)(3x+6)$

"FOIL"

Figure 5.1 Multiplication of two binomials from Sidney's lessons (Day 2/3).

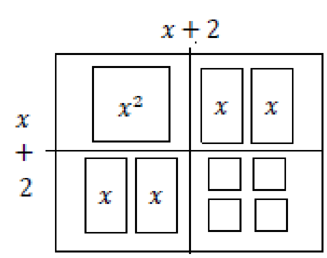
When Sidney originally submitted his Task Log Sheets, he listed each lesson that contained a polynomial property (with related practice problems) collectively as one "task" and label them as "reasoning-and-proving" or "maybe" or "some" reasoning-and-proving. During his second interview (in person), Sidney was asked to rate the reasoning-and-proving potential of each individual exercise. At that time he labeled each exercise in the lesson shown in Figure 5.1 as non-reasoning-and-proving, which matched the rating of the coders. In retrospect, however, while these exercises do not fall under one of Thompson, Senk, and Johnson's *exercise* codes, the exercises in the lesson shown in Figure 5.1 do try to justify a property for multiplying binomials in the manner of Thompson, Senk, and Johnson's *narrative* codes. In this case, because the lesson is so scaffolded, there is little opportunity for students to engage in reasoning, but the property is justified with a specific case using an area model and the distribution method.

Figure 5.2 shows Sidney's factoring quadratic trinomials lesson (the italicized words are Sidney's answers). There is only one example and a suggested drawing before students are asked to write a conjecture about factoring. It feels more like a review than a novel concept for students. However, it is possible that Sidney led a rich discussion about how to write a general statement for this property and justify it with a deductive argument. Without observational data, we have no evidence either way.

General form of a trinomial: $x^2 + bx + c$ coefficient is 1

With algebra tiles, demonstrate the following expression: $x^2 + 4x + 4$

Draw your expression here:



Now, your goal is to work backwards to get the two binomials that form this trinomial. Go ahead and using the picture you just drew, write the two binomials along the sides of the model.

Write a conjecture that describes about [sic] how to return back to the original two binomials.

You want the sum to equal the middle term and the product of these numbers equal constant

What did we come up with?

Two key aspects when we *factor* a trinomial?

1. *Sum of the numbers equals the middle term*
2. *Product equals the constant*

What is the *factor* form of $r^2 + 11r + 24$?

Figure 5.2 Factoring a perfect square trinomial from Sidney's Lessons (Day 11).

Because the research on the availability of reasoning-and-proving tasks in textbooks clearly indicates the limitation of this resource (e.g., Thompson, Senk, & Johnson, 2012), the CORP course included a chapter on modifying tasks to further equip participants to offer reasoning-and-proving opportunities to their students. The participants looked at several versions of tasks—unmodified textbook tasks and modified versions—and created a strategy list to help them modify tasks in the field. The strategies included engaging students in investigations and conjectures instead of just giving answers, provide all students with access to a task by making initial observations, require students to provide a mathematical argument or proof, remove unnecessary scaffolding, and ask students to generate empirical examples and look for patterns. The participants in the study reported herein, though, modified very few exercises to improve their potential to engage students in reasoning-and-proving (only 7 exercises out of the 1,453 exercises selected by the participants).

Karen's modifications had a neutral effect on the reasoning-and-proving exercises in her lessons. Uma inserted two conjectures into lessons (see Figure 4.14 and Figure 4.20) and included an investigate-a-claim question from a previous chapter (with a neutral effect) into her chapter review packet. In addition, though, Uma created a disproof and a proof question for her classroom artifact packets and asked her students to present and justify their work from her second classroom artifact packets (see Figure 4.21). If one combines the modified exercises with the exercises Uma created, Uma used many of the strategies discussed in the CORP university course. The only strategy she reversed was "remove scaffolding." Like Jonathan, Uma at times inserted more scaffolding into a problem than was in the original exercise. Jonathan cited the ability level of his students for the reason he had the scaffolding; it was assumed Uma had the same reason. Like Uma, Jonathan inserted two conjectures into lessons (see Figure 4.37). In looking at all of his exercises and modifications, Jonathan appeared to focus on the strategy of providing all students with access to the task by first making observations about a situation before moving on to more focused work. Finally, Sidney inserted three conjectures into his lessons on polynomial operations (e.g., Figure 4.30), but he did not engage his students in much exploration (sometimes only a single example) before the students made their conjectures. There is some evidence (in the similarity of students' answers) that the students had seen this content already.

During her interview, Uma expressed a desire for a book of proof tasks to use as a resource. The data shows that none of the participants used any proof activity from the CORP course, nor did they make much use of the internet or other books as ancillary resources. The tasks in the CORP course included 6 pattern tasks, 6 number theory tasks, 4 geometry tasks (1 linear pair/vertical angles, 1 parallel line proof, 1 Pythagorean identify, and one constructing a parallelogram), and 2 algebra tasks (graphing a line with intercepts and A Sticky Gum problem). Only Karen's curriculum overlapped with an activity in the CORP course during her 15-day data collection period, and there were already similar proof tasks in her textbook. That being said, Thompson, Senk, and Johnson (2012) suggest that it is useful for students to look at examples and nonexamples, correcting mistakes, and evaluating arguments as they develop their ability to prove. A teacher could take a traditional curriculum and modify some existing exercises to include these types of reasoning activities. However, there is no evidence that the four participating teachers in the current study did so.

5.2.2 Implementation of reasoning-and-proving tasks

The results of this study suggest that the participants were able to implement cognitively-demanding reasoning-and-proving tasks at a high level. Throughout the CORP university course, participants read and discussed narrative cases which described teachers engaging their students in varying degrees of reasoning-and-proving. The cases illustrated the challenges teachers face in their own classrooms. The teachers in the cases supported the capacity of their students to reason-and-prove by the tasks, tools, and talk used in their classrooms. In particular, the pre-service teachers examined how the narrative case teachers' instructional decisions helped or hindered their ability to implement high-level reasoning-and-proving tasks. The finding that the participants in this study were able to maintain the high level of cognitive demand of their reasoning-and-proving tasks between potential and implementation is significant because high-level tasks usually decline during implementation (Stein, Grover, & Henningsen, 1996; Stigler & Hiebert, 2004).

According to the IQA rubrics for potential and implementation of tasks, a high-level task asks students to make their reasoning explicit or implicit. As such, reasoning-and-proof tasks that ask students to identify patterns and form generalizations on those patterns, make or investigate conjectures and support conclusions with mathematical evidence and/or create a proof or find a counterexample, or evaluate an argument or explain how to outline an argument of a particular type are cognitively-demanding tasks. The participants of the study described herein asked students to create proofs, find counterexamples, identify patterns and make generalizations and conjectures.

Only Karen selected three proof tasks to implement and for which to collect classroom artifact packets; the other participants selected 2 proof tasks (Uma) and no proof tasks (Sidney and Jonathan, although Sidney thought he was implementing proof tasks). Asking students to make a conjecture is still considered to be a high-level task, however, and making conjectures accounted for all of Sidney's and Jonathan's tasks in their student work samples. Thus, all of the participants selected tasks for their classroom artifact packets that were cognitively demanding.

Most of the participants maintained the level of cognitive demand between the potential of the tasks and the implementation. Karen asked for proofs and got them. The question with Karen's students' work is whether or not students were making their reasoning explicit with two-column proofs or whether they were merely mimicking a procedure. Uma used a creative assortment of tasks (disproof, procedure, and proof), and her students' work suggests that they understood how to disprove a statement (by finding a counterexample) but had trouble proving a statement (animal pen task). Uma reported that she had a conversation with her students about why a circular pen enclosed the maximum area, but without student interview data, it is impossible to know what each student learned about proof from the conversation (more will be said about this later). Sidney's tasks seemed to be a little inappropriate for the level of some of his students, which might account for the decline of the high-level cognitive demands of his tasks (Stein, Grover, & Henningsen, 1996). Finally, Jonathan asked for conjectures and got them from his students, although the amount of scaffolding in his tasks prevented the tasks from having the highest-level of cognitive demand.

Uma presents an interesting implementation case in her third student work sample (find the shape of an animal pen that will maximize the area enclosed if the perimeter is fixed). She labeled her first student work sample (Uma 3-1) as "exceeding expectations" but also "nonproof" because the student used an empirical argument (see Figure 4.27). Uma 3-2 was labeled "exceeding expectations" and proof, because the student provided a reason for why a circular-shaped pen would enclose the maximum area ("The circle is bigger because whatever shape you make it basically stretches the perimeter to the max because it has no corners."). Uma 3-3 had a typical response, which was to calculate the areas of a square and two rectangles (see Figure 4.26). This student received a rating from Uma of "met expectations" and "nonproof." On her classroom artifact packet Cover Sheet for her third task, Uma wrote:

> Many students were stuck with only using squares and rectangles; we had much more conversation than anticipated. I was impressed with how we discussed the task and all the possible shapes. I wish I could have been more organized in conversation and had students record all answers.

Without observational data, it is difficult to tell whether Uma held all students to the standard of proving that a circle would enclose the maximum area (by a method appropriate to her level of students) or not. The way she rated her expectations of the student work indicates that she did not, but her comments about the implementation imply that she did. It could be that proving that a circle was the best shape did not matter to Uma. Knuth (2002b) found that some practicing teachers recognized informal proofs (based on empirical arguments) as invalid but acceptable stepping stones on the way to proof. However, Chazan and Lueke (2009) found that when teachers tacitly accept empirical arguments, their students' conception of valid proofs include empirical arguments.

5.2.3 Evaluation of reasoning-and-proving products

The results of this study suggest that participants were able to articulate and use criteria for proof that included most if not all of the core elements of proof. In solving a sequence of tasks described by G. J. Stylianides (2009) and examining samples of student work for the task "prove that the sum of two odd numbers is an even number," participants reviewed the limits of empirical arguments and constructed a list of the core and auxiliary elements in mathematical proof. Research has shown that an "empirical proof scheme" is common among students and teachers (e.g., Chazan, 1993; Harel & Rabin, 2010; Knuth, 2002b; Stylianides, G. J. & Stylianides, A. J., 2009) and that teachers have difficulty making decisions about rigor when engaging students in reasoning-and-proving (Knuth, 2002b, 2002c; Stylianides, A. J., 2007). The fact that the participants in this study included most of the core elements in their definition of proof and in their rubrics is significant.

From the discussion in the CORP course regarding what constitutes reasoning-and-proving for secondary students, the participants generated a list of core elements of proof (with guidance from the course instructor who had A. J. Stylianides' 2007 definition of proof, see Figure 5.3). When asked for their definitions of proof, each participant included that the statements and definitions that area used in the argument must be ones that are true and accepted by the community because they have been previously justified. This was not a surprise, because this core element of proof ("accepted by the community") was something new for the participants. This core element resonated with the participants as they imagined teaching proof to students. All of the participants stated that a proof shows a claim is true or not true, but not everyone addressed the domain under which the statement is true. Only one participant—

Karen—included logic in her definition of proof. Of all the core elements, "logical flow" got the least explicit attention in the CORP course. It is possible that an apparent lack of attention to the logical flow of an argument accounts for the participants not modifying or creating tasks to correct a mistake or evaluate an argument.

Core Elements

- States defined terms and previous proven properties with respect to the audience
  Defines all terms (unless accepted by audience) and states previously proven properties
- Every step is supported mathematically without error
- Organized in a logical flow
- Reaches and states a supported conclusion
- Clearly generalized for a specific domain

Auxiliary Elements

- providing examples that explain
- inclusion of ALL definitions, properties, etc. regardless of whether the class has already proven them or not
- form of the proof (picture, algebraic, words, etc.)
- explanatory power
- accessible for all audiences
- type of proof (induction, contradiction, etc.)

Figure 5.3 Criteria for a valid proof as developed by Sidney's and Jonathan's CORP class on May 22, 2012.

The only measure explicitly used to evaluate student work samples of proofs in the CORP university course was the core elements of proof. In Karen's rubrics for her student work samples, she included "formal proof" in her expectations, which she defined roughly as a detailed, understandable argument that uses statements and theorems with justification to arrive at a goal. Based on her rubrics, Karen was focused on justification. In contrast, Uma's rubrics for her student work samples seemed to be more focused on correctness of the answer with explanations and pictures that were convincing. Uma did not specify who the argument should convince (the teacher? other students?) in her written rubrics. While Uma included the auxiliary element "pictures" in her rubric, it is likely that Uma was suggesting a way for students to support their arguments.

Jonathan did not assign any proof tasks to his students during the 15-day data collection period, but it is worth noting that his rubrics focused on the quality of the students' conjectures and included "makes generalized, detailed statements." He also expected his students to use correct mathematical vocabulary (a help when communicating). Jonathan stated that he was laying the ground work for engaging his students in proof next year; his attention to generalized, detailed, correct mathematical conjectures implies an awareness of the core elements of proof.

In communicating expectations, no participant developed a list of the core elements of proof with his or her students. Consequently, students had no reference material to help them construct proofs. The only participant who came close was Karen, who had her students keep a tan composition book with definitions and theorems. Informally, however, all of the participants held their students to justifying their mathematical statements in class, and Karen's and Jonathan's students had many examples of graded work to use as models of expectations.

5.3 REASONS FOR THE RESULTS

While the results of this study are encouraging, the study does raise some questions as to why there was inconsistency among the participants regarding the extent to which they were able to select, implement, and evaluate the products of reasoning-and-proving tasks. The CORP course likely changed the participants in some way, but the fact that the participants came into the course with different skill sets, abilities, and understandings suggests that the participants were not likely to leave the course with the same skills set, abilities, and understandings. Since the participants were preservice teachers during the CORP course, it was not possible to assess their ability to engage their own students in reasoning-and-proving before the course. This research project did not assess the mathematical ability of the participants before or after the course, either. It is not the intention of the CORP course to help participants fine tune their understanding of logic, explore common proof techniques in specific mathematical domains, or develop their classroom management techniques. The early-career teachers who participated in this study are—like all teachers—on a continuum of skills and beliefs. The research discussed in Chapter 2 listed numerous studies which revealed that in general, teachers in the United States do not engage their students in reasoning-and-proving, and when they do, the work tends to be formulaic, limited in scope, and offers few opportunities to develop understanding. Regardless of the personal understandings of mathematics or ability to teach, each participant in the study described herein engaged their students in some reasoning-and-proving learning opportunities, even if the level of potential and implementation was unequal.

Another strong possibility is that the constraints of the school systems in which the participants work had an overwhelming influence on the extent to which the participants engaged their students in reasoning-and-proving. Table 5.1 lists the perceived constraints discussed during the interviews with the participants, and Table 5.2 captures the perceived facilitating elements the participants' districts offer. In looking at the constraints data, it appears that traditional curriculum, daily pacing guides and state and district testing hamper the extent to which participants engaged their students in reasoning-and-proving. This is not to suggest, however, that every constraint and facilitator has equal influence. For instance, both Karen and Sidney are partially evaluated as teachers based on their students' test scores (see Table 5.3). Therefore, Karen and Sidney have to take the state and district exams seriously if they want to keep teaching in their districts. Also, while Karen was using a reform curriculum which devoted an entire unit on proof (and congruence), the curriculum took a quasi-traditional approach with respect to the types of statements students were asked to prove and to the containment of formal proof to one unit.

Table 5.1 Constraints Faced by the Participants

Constraint	**Karen**	**Uma**	**Sidney**	**Jonathan**
Years of experience			X	X
Below-grade level students		X		X
Instructional Time	X		X	
Professional Development Time				
Beliefs and attitudes of other teachers		X	(X)	
Behavior Issues				X
Traditional curriculum	(X)	X	X	X
Scripted curriculum (daily pacing guide)	X	X	X	
State and District Testing	X	X	X	
Learning support	X			
Algebra 1 vs. Geometry			X	X

Table 5.2 Facilitating Elements for Engaging Students in Reasoning-and-proving

Facilitators	Karen	Uma	Sidney	Jonathan
Years of Experience	X	X		
At or above grade-level students	X		X	
Instructional Time		X		
Professional Development Time	X		X	
Collaborative work with department members	X		(X)	
Curriculum promotes students' reasoning	X			
Freedom to Design Curriculum				X
CCSSM (adopted)	X			X
Geometry vs. Algebra 1	X	X		

Table 5.3 Evaluative Measures for Participants by Their School Districts

Evaluative Measure	Karen	Uma	Sidney	Jonathan
Observations (with pre-and post-conference)	X		X	
Observations (walk-through)	X	X	X	X
Student Test Scores	X		X	
Student Perception Surveys	X	X		
Lesson Plans	?			X
Teacher portfolio (self-improvement plan)			X	

Notice that working with other teachers is listed under both constraints and facilitators. Karen shared that her department members are "pretty good about getting together and sharing ideas" (personal communication, February 20, 2013). Uma, on the other hand, chafed against the chapter exams her traditional, fellow teachers constructed. She stated, "if people would actually teach their students for the way that they learn, we wouldn't have [poor SOL exam scores]" (personal communication, February 24, 2013). Sidney was more neutral about the influence of his fellow teachers. On one hand, he was grateful for their support and shared resources (which he was strongly encouraged to use to keep consistency in the department), but on the other hand, they did not give much attention to reasoning-and-proving. Jonathan had no one to help or hinder him as he tried to engage his students in reasoning-and-proving.

There might be other constraints and facilitators present for each participant of which they were unaware or did not vocalize in the interviews. For instance, Sidney struggled in his university mathematics courses and with the mathematics in his methods courses; his understanding of mathematics probably influenced his ability to enact reasoning-and-proving. Another unvoiced constraint or facilitation might be classroom norms. No participant volunteered that by the time data was collected, negative student behavior diverted participant's attention away from reasoning-and-proving or that the participants' beliefs about reasoning-and-proving influenced their desire to offer reasoning-and-proving learning opportunities with the students. While the CORP course did not simply *tell* participants that reasoning-and-proving was important, it is unclear the extent to which the course activities situated reasoning-and-proving in the participants' internal hierarchy of important mathematical practices and content for students. These unvoiced or unrecognized constraints and faciliators could have influenced the participators' decision-making with respect to reasoning-and-proving.

Another question raised by this study is why the participants did not modify more exercises to increase the exercises' reasoning-and-proving potential. Did the CORP course offer insufficient training? According to several lines of research, transformative professional development that can reform teaching should include opportunities for teachers to study the big ideas of the discipline (Loucks-Horsley & Matsumoto, 1999), develop flexible understandings of the mathematics they will teach (Ball & Cohen, 1999; Thompson & Thompson, 1996), practice making complex and subtle decisions about teaching (Little, 1993), learn self-monitoring and analysis (e.g., Stigler & Hiebert, 1999), and do this external to real classrooms and in real time (Ball & Cohen, 1999). The CORP course contained all of these elements. The participants studied the importance of proof and the core elements, how to select, modify, and implement a reasoning-and-proving task, and how to evaluate an argument. Participants discussed using tasks, tools, and talk to enact reasoning-and-proving. The participants also used samples of authentic practice, such as classroom tasks, samples of student work, and cases of teaching episodes. These samples of authentic practice were intended to help teachers discuss abstract and complex ideas, the dilemmas of teaching, and reveal students' thinking about mathematics (Smith, 2001). Finally, the participants studied reflection and practiced presenting a few solutions which helped preservice teachers develop core practices (Grossman et al., 2009b). Participants did not spend much time rehearsing orchestrating discussions, though, or rehearsing conveying the importance of reasoning-and-proving as a core mathematical practice. According to the results of research on successful professional development for preservice teachers, the CORP course should have prepared teachers to enact reasoning-and-proving tasks in their classrooms, but perhaps the course could have included additional opportunities to enact teaching reasoning-and-proving. The study described herein does suggest that the participants of the CORP course did indeed select, implement, and evaluate the products of reasoning-and-proving tasks far more than most teachers do (Porter, 1993). It is possible, though, that early career teachers need more, continuing, and differentiated support to apply the knowledge they learned in the CORP course.

5.4 CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

During their university pre-service training, the four participants in this study learned how important reasoning-and-proving was for every student in mathematics education; an entire course was devoted to this topic and it is in the major standards (NGA Center & CCSSM, 2010; NCTM, 1989, 2000). Much time was devoted in the course to the core elements of proof. Once the participants were in the field, most of the participants engaged their students in more reasoning than formal proving. Each participant stated that they “push kids to justify” (Karen, personal communication, February 20, 2013), back up statements, explain the origin of ideas, and share why statements are true. As previously mentioned, teachers might need more support during the school year as they engage their students in reasoning-and-proving. Future iterations of the CORP course could be offered during the school year rather than in the summer so that teachers could continually make connections to their practice.

In the winter, Sidney’s department chair sent him to a district professional development meeting on reasoning-and-proving because he did not think anyone else would want to go. Sidney shared the thinking of the other teachers in the room (personal communication, February 26, 2013):

> So one thing we talked about…we can do proof and reasoning in our classes but if they aren’t going to be using it next year or using it in high school there’s no point, and they

> don't have any experience in high school, they're not going to be used to it and it's going to cause more problems again.

These comments point to a need for systematic attention to reasoning-and-proving throughout the mathematics education curriculum, K-12. Recently, one large urban district addressed this need with a pilot of all of their practicing secondary teachers. The teachers attended monthly sessions on reasoning-and-proving then returned to their classrooms to try what they learned. This model merits attention and could be studied.

The insufficient attention to reasoning-and-proving in textbooks and the participant's minimal exercise modifications point to a need for more reasoning-and-proving resources for participants. If such resources are already available, then the participants are unaware of the resources' existence or the participants did not have the time or inclination to find them. It is also possible that the available resources on the market or on the internet are not tied closely enough to a participants' curriculum to make the resources convenient and efficient choices, or that there is not enough support attached to the resources to make then an attractive option. Teachers need more resources that are educative and provide teachers with convenient and relevant materials to use as well as support for using them.

Research suggests that many teachers do not understand key aspects of reasoning-and-proving or why it matters (Harel & Rabin, 2010; Knuth, 2002c; Porter, 1993). This study—while it represents a very narrow slice of the practice and understanding of participants—provided an analysis of four trained teachers that makes salient that while there are individual differences, the teachers appear to understand why reasoning-and-proving is important, can articulate the core elements of proof, and can identify and implement tasks that get at the key aspects of reasoning-and-proving. Clearly additional work is needed to understand the level of support teachers need and what form that support should take to be maximally effective. Follow–up studies could paint a more detailed picture of the participants' content knowledge, beliefs about reasoning-and-proving, and classroom norms. Subsequent research could study the impact of the CORP materials on established, practicing teachers, and collect initial information (content knowledge, reasoning-and-proving understanding and beliefs, and classroom norms), data on enactment during the professional development (observations, classroom artifact packets, student interest surveys and formative assessments), and post-professional development information (content knowledge and understanding of reasoning-and-proof), in order to get a more complete picture. This additional information could help researchers determine which elements of the CORP materials have an impact on specific aspects of practice.

Appendix A

EMPIRICAL STUDIES OF REASONING-AND-PROVING

Article	**Date**	**subjects**	**Major Findings**
Alcock & Weber	2005	students, undergraduate, British	Students had difficulty with assessing the logic of an entire argument, not just the last line. Statements in an argument need to be warranted from previous statements.
Bell	1976	students, high school, British	Students had difficulty with using precise definitions
Bieda	2010	teachers, middle school	Teachers' feedback to students was insufficient to establish standards of mathematical proof.

Chazan	1993	students, high school, geometry	Students did not trust arguments based on deductive reasoning so preferred empirical reasoning, or blended the two methods while proving. Students still thought counterexamples were still possible even after a concept was proven. In addition, Chazan found that a person's beliefs about a mathematical object may affect their perception of the truth of a proof's conclusion based on that object.
Coe & Ruthven	1994	students, undergraduate, successful	Students had trouble verbalizing their thoughts and abstracting concepts; they also abandoned the context of the task and preferred empirical arguments.
Doyle	1988	teachers, middle school	Teachers find classroom management difficult when students work on novel tasks, which sometimes led to a smoothing of the curriculum.
Edwards & Ward	2004	students, undergraduate, math majors	Students have difficulty with definitions; authors make suggestions about coordinating definitions and theorems in proofs.
Edwards	1999	students, high school, algebra 1	A student will only seek justification for a claim if he believes the statement to be true in general. Students prefer empirical justification and look to the teacher as an external authority. Students were able to generate counterexamples for false statements. Edwards suggests eliciting students' own formal arguments and justifications and using those for class discussion.

Ellis	2011	teachers, middle school	Teachers had difficulty enacting reasoning-and-proving tasks that required generalization. Ellis suggests prompting students to predict the outcome of an experiment or hypothesis.
Fischbein	1982	students, high school	Students have a difficult time with the concept of proof, specifically with respect to generality (only 24.5% of the students accepted that no further checks were necessary after proving a claim). Fischbein also found that students' beliefs about mathematical objects may affect the perception of truth of a proof's conclusion.
Fischbein	1999	students, high school	When students' intuition is in conflict with a formal proof, make sure the formal proof wins. Fischbein recommends pushing students to prove even apparently trivial properties because properties are not automatically applicable to every math operation.
Galbraith	1981	students, high school, Australian	One-third of students did not understand the role of counterexamples in refuting general statements; students also had difficulty with chaining inferences in deductive arguments, generality, and external authority.
Goetting	1995	students, undergraduate, preservice	80% of the preservice teachers studied preferred empirical arguments.
Hadas, Hershkowitz, & Schwarz	2000	students, high school, Israeli	Students found empirical arguments (created with Dynamical Geometry Environments) convincing because they believed in technology over reason, however, the students did use deductive reasoning to explain contradictions.

Harel & Rabin	2010	teachers, high school	Teachers favored empirical arguments, employed an authoritative proof scheme, and used the instructional sequence: present a rule, provide an illustration, and provide examples. To combat authoritative proof schemes, Harel & Rabin suggest that the teacher responds to students' questions by probing their reasoning or asking them to check correctness for themselves. The teacher should allow students to debate and resolve disagreements and the teacher could offer deductive justifications.
Healy & Hoyles	2000	students, high school, successful	Students had difficulty constructing proofs and they favored empirical arguments. Students perceived that teachers favored formality over communication. Students were able to identify valid proofs, knew that empirical arguments were limited, and understood generality. When students believed in the truth of a statement, they accepted empirical arguments as sufficient.
Healy & Hoyles	2001	students, high school, successful	Students linked properties--and their place in proofs--with constructions.
Inglis & Alcock	2012	students, undergraduate	Students focused on format rather than checking warrants for assessing the validity of an argument.

Knuth	2002c	teachers, high school	4/17 teachers thought the role of proof was to display students' thinking, and 7/17 thought the role of proof was to explain why a statement was true. 13/17 teachers thought proof helped students develop logical thinking skills. 10/17 considered proof a social act, used to communicate mathematical thinking. The teachers thought that proof was a logical or deductive argument that demonstrated the truth of a claim. However, 14/17 teachers thought that proof was not appropriate for all students, and 17/17 teachers thought that even though empirical arguments are not valid proofs, empirical arguments were fine for lower-level students. 4/17 teachers tied "formal" proofs to the two-column format. The type of curriculum (reform or traditional) or course-load had no effect on the teachers' answers. The teachers interpreted the curriculum author's use of "proof" to fit the teacher's conception of proof.
Knuth	2002b	teachers, high school	Teachers were better able to identify correct proofs than incorrect proofs (93% of teachers correctly identified valid proofs. One-third of non-proofs (some empirical) were as proofs.) Teachers thought that proof is a social act, that the role of proof was to explain procedures, not to promote conceptual understanding. Teachers had difficulty with generalization (similar to Fischbein and Martin & Harel). If a teacher recognized a particular method he considers as valid (even if the teacher did not understand the method but the mechanics of the argument were correct) or recognized sound mathematics, the teacher found the posed proof as convincing. Teachers used mathematically grounded criteria for accepting an argument as proof while using qualitatively grounded criteria for making distinctions among proofs.

Knuth & Sutherland	2004	students, middle school	40% of students favored empirical arguments over deductive arguments; 30% selected deductive arguments over empirical arguments. Some students had difficulty with generality and did not see the difference between proof-by-exhaustion and empirical arguments, but other students produced and selected general arguments, recognized the limitation of empirical arguments, and correctly used proof-by-exhaustion. Question: do students rely on empirical arguments because they are unable to produce a general proof? Healy and Hoyles 2000 made a similar observation.
Ko & Knuth	2009	students, undergraduate, Taiwanese	Students had difficulty with coordinating mathematical knowledge (e.g., continuity with limits of functions), counterexamples, and producing proofs. Students also manipulated symbols without understanding and did not support warrants.
Küchemann & Hoyles	2001	Students, high school, successful	Students may be influenced by teaching and textbooks; students had difficulty with using context and generating patterns and seemed to use their perceptions rather than theorems in writing proofs.
Laborde	2000	students, high school	Dynamic geometry environment fostered interaction between construction and proof, between doing on the computer and justifying by means of theoretical arguments.
Mariotti	2000	students, high school, geometry, Italian	Students were able to move from empirical to formal justifications by the organization of tasks and dynamic geometry environment.
Mariotti	2001	students, high school, geometry, Italian	Students developed an axiomatic system by using technology to build constructions, collectively revised the constructions during class discussion, and then the relationship between axioms, definitions, and theorems and use of tools emerged.

Marrades & Guttierez	2000	students, high school, geometry, Spanish	Students progressed from empirical to formal justifications over 30 weeks; this was helped by the organization of tasks and a dynamic geometry environment; notebooks were kept for record of theorems, etc.
Martin & Harel	1989	students, undergraduate, preservice mathematics, elementary	Students thought empirical proofs were valid.
Martin, McCrone, Bower & Dindyal	2005	students, high school, geometry	Students had difficulty with generality. Authors suggested posing open-ended tasks and engage students in verbal reasoning.
McCrone & Martin	2009	students, high school	Students no concept that building an axiomatic system was one of the reasons for proof; students thought that empirical arguments were fine for students but not preferred by teachers. While students thought a logical flow to the argument was important, they focused on checking the format rather than checking warrants.
Mingus & Grassl	1999	students, undergraduate, preservice mathematics teachers	Preservice elementary teachers thought that proof was to demonstrate or confirm relationships and proof made students question why things work the way they do. The preservice elementary teachers said that high school geometry was the last time they saw proof. Preservice secondary teachers thought proof was explanatory, logical, and convincing, and the role of proof was to maintain and advance the structure of mathematics. Only 17% of preservice elementary teachers thought proof should be taught in grades K-6; 63% of preservice secondary teachers thought proof should be taught in grades K-6.

Moore	1994	students, undergraduate	Students had difficulty with the concept of proof, understanding definitions, language, and notation, and they had trouble launching into a proof. Moore suggested that teachers generate and use examples, apply definitions, and use definitions to structure proofs.
Morris	2002	students, undergraduate, preservice mathematics, elementary and middle school	40% of the students accepted at least one inductive argument and 47% thought exhaustive checks were necessary regardless of the proof (e.g., problems with empirical arguments and generality). On the other hand, 47% of the students rejected inductive arguments and 30% said exhaustive checks were unnecessary after proof. Students recognized that mathematics is not infallible (historical reasons) and any axiomatic system is based on primitives (definitions, postulates which may or may not be true)…but no one gave this as a reason to mistrust generality. What a person understood/believed about the premise effected what type of argument they considered sufficient and what they trusted.
Porteous	1990	students, middle and high school	Students favored (but somewhat mistrusted) empirical arguments and had difficulty with generality. 48/50 students correctly used counterexamples. When presented with a particular case, over half empirically checked it rather than appealing to the proof of the general case that they had previously been shown and had presumably accepted (note: 75% of the students accepted the premise as true from the beginning).
Porter	1993	teachers	There is little time to teach proof and not much gets taught.
Raman	2003	students, undergraduate	Students had difficulty accessing relevant knowledge and relied on heuristic ideas while constructing proofs. Arguments grounded in empirical data gave sense of understanding but not conviction.
Schoenfeld	1986	students, high school, geometry	Students had difficulty with generality (made conjectures that contradicted already proven statements).

Schoenfeld	1989	students, high school, geometry	Students had difficulty with the concept of proof and preferred empirical arguments. Students contradicted the conclusion of a proof they had constructed.
Selden & Selden	2003, August	students, undergraduate, math majors	Students exhibited the following difficulties in proof-making: generalization, use of theorems, notation and symbols, nature of proof, and quantification. Students had a static view of mathematics (correctness of an answer depends entirely on selecting the right algorithm and on implementing its steps correctly; knowing algorithms means mathematical competence)--this view made constructing proofs difficult. Students also had difficulty with conservation of relationships and thought real number laws were universal. Students began proofs with the conclusion, left holes in their reasoning, and employed circular reasoning. At times their proofs were locally unintelligible (format was fine and correct symbols were used but the assertions were incomprehensible or incorrect). Sometimes students proved a weakened version of a theorem, and sometimes the students failed to notice restrictions on variables. Students also substituted with abandon and at times viewed non-real statements (e.g., cos x = 3) as existing because the statement was written. Students also overextended symbols and failed to adapt notation from one context to another. Students used information out of context. At times students used the converse of theorems without verifying its validity (e.g., used "if, then" when "if and only if" is meant).
Selden & Selden	2003	students, undergraduate, preservice mathematics, secondary	Students had difficulty with validating proofs: students attended to global/structural errors, such as proving the converse of the statement, attended to frameworks, and reflection during interviews.

Senk	1985	students, high school, geometry	Students used circular reasoning and had difficulty launching proofs, adding auxiliary lines to diagrams, and making multiple deductions. Students had a 50-77% success rate with fill-in-the-blank proofs and proofs with figures. Senk suggested teaching the meaning of proof, teach when, why, and how to transform a diagram, and teach chains of deductive reasoning.
Simon & Blume	1996	students, undergraduate, preservice mathematics, elementary	Students had difficulty with understanding proof as a means to understand mathematics. Students preferred empirical arguments and questioned who has the authority to validate a proof.
Stylianides, A. J., & Stylianides, G. J.	2009	students, undergraduate, preservice mathematics, elementary	Students recognized that empirical arguments were insufficient but still used them.
Stylianides, A. J., Stylianides, G. J. , & Philippou	2004	students, undergraduate, preservice mathematics, Cyprus	Preservice elementary students had difficulty with contraposition equivalence rule in symbols but were fine if the statement was made verbally (if p = q, then not q = not p) and used format as the method to judge the validity of a proof. Preservice secondary students did not have difficulty with the contraposition equivalence rule in symbols or verbal.
Stylianides, G. J., Stylianides, A. J. , & Philippou	2007	students, undergraduate, preservice mathematics, Cyprus	Doing: trouble with mathematical induction
Tabach, et al.	2011	teachers, secondary, Israeli	Teachers over-value generality of symbolic mode of representation and under-value verbal modes of representation.

Tinto	1990	students, high school, geometry	Theorems, definitions, axioms, and geometry had no significance as an axiomatic system for students. The goal of textbook exercises of proofs appears to be efficiency.
Varghese	2009	students, undergraduate, preservice mathematics, secondary	Students thought proof only applied to certain content areas and was only appropriate for the best students. Students thought proof was for verifying something already known (9 out of 17 students), derivation (2/17), logical argument (2/17) and justification (2/17). Students thought the best way to teach proof was step-by-step demonstration (13/17 students), constructivist approach (3/17), and social interaction (1/17). Students recognized that their discomfort with proof will impact their students' opportunities to learn proof.
Weber	2001	students, undergraduate, math majors	Students lacked strategic knowledge of a domain's proof techniques, choosing theorems, and applying theorems; students had difficulty accessing relevant knowledge.
Williams	1979	students, high school	Students had difficulty with the concept of proof, definitions, empirical arguments, indirect proof, and logically equivalent statements. Less than 20% of students understood indirect proofs. Students questioned why they had to prove intuitively obvious statements. 20% of students did not realize what a general proof meant, but 31% did understand.

Appendix B

SUMMARY OF EMPIRICAL STUDIES ON REASONING-AND-PROVING BY AREA OF DIFFICULTY

Subjects	**Students: Secondary**		**Students: Undergraduate**				**Teachers**	**Empirical Studies**
Area of Difficulty	**6-8**	**9-12**	**Non-specific**	**Math major**	**Preservice elementary**	**Preservice secondary**	**Secondary**	
Audience (appropriate for?)					X	X	X	Knuth (2002c), Mingus & Grassl (1999), Varghese (2007)
Concept of Proof		X	X	X			X	Fischbein (1982), Knuth (2002c), Moore (1994), Selden & Selden (2003), Williams (1979)
Context		X	X					Coe & Ruthven (1994), Küchemann & Hoyles (2001)
Coordinate knowledge			X					Ko & Knuth (2009)

Subjects	Students: Secondary		Students: Undergraduate				Teachers	Empirical Studies
Area of Difficulty	**6-8**	**9-12**	**Non-specific**	**Math major**	**Area of Difficulty**	**6-8**	**9-12**	**Non-specific**
Definitions		X	X	X				Bell (1976), Edwards & Ward (2004), Moore (1994), Williams (1979)
Diagrams		X						Senk (1985)
Empirical Arguments	X	X	X		X	X	X	Chazan (1993), Coe & Ruthven (1994), Edwards, L. D. (1999), Goetting (1995), Hadas, Hershkowitz & Schwartz (2000), Harel & Rabin (2010), Knuth & Sutherland (2004), Martin & Harel (1989), Morris (2002), Schoenfeld (1989), Simon & Blume (1996), Stylianides & Stylianides (2009), Williams (1979)
Format							X	Knuth (2002c), Tabach et al. (2012)
Language			X					Moore (1994)
Launching		X	X					Moore (1994), Senk (1985)
Logic (chaining inferences)		X						Galbraith (1981), Selden & Selden (2003), Senk (1985)

Subjects	Students: Secondary		Students: Undergraduate				Teachers	Empirical Studies
Area of Difficulty	**6-8**	**9-12**	**Non-specific**	**Math major**	**Area of Difficulty**	**6-8**	**9-12**	**Non-specific**
Logic (lack understanding)		X	X	X				Alcock & Weber (2005), Hadas, Hershkowitz & Schwarz (2000), Ko & Knuth (2009), Selden & Selden (2003), Williams (1979)
Nature of Proof				X				Selden & Selden (2003), Weber (2001)
Notation and Symbols			X	X			X	Moore (1994), Selden & Selden (2003), Tabach et al. (2011)
Pedagogy						X	X	Bieda (2010), Doyle (1988), Harel & Rabin (2010), Porter (1993), Varghese (2007)
Rigor	X	X					X	Fischbein (1982), Healy & Hoyles (2000), Knuth (2002b, 2002c), Küchemann & Hoyles (2002), McCrone & Martin (2009)
Role of Proof (to understand)		X			X		X	Knuth (2002b), Mingus & Grassl (1999), Simon & Blume (1996), Tinto (1990)
Role of Proof (to create an axiomatic system)		X						McCrone & Martin (2009), Tinto (1990)

Subjects	Students: Secondary		Students: Undergraduate				Teachers	Empirical Studies
Area of Difficulty	**6-8**	**9-12**	**Non-specific**	**Math major**	**Area of Difficulty**	**6-8**	**9-12**	**Non-specific**
Technique (choosing, using, and misusing theorems and converses)				X				Selden & Selden (2003), Weber (2001)
Technique (types of proof)		X	X	X	X	X		Chazan (1993), Knuth & Sutherland (2004), Ko & Knuth (2009), Stylianides, Stylianides & Philippou (2004, 2007), Weber (2001), Williams (1979)
Validation				X			X	Knuth (2002b), Selden & Selden (2003)
Validation (based on format)		X	X		X			Inglis & Alcock (2012), McCrone & Martin (2009), Stylianides, Stylianides & Philippou (2004)
Validation (authority?)		X			X			Edwards, L. D. (1999), Galbraith (1981), Simon & Blume (1996)

Appendix C

SUMMARY OF THOMPSON, SENK, AND JOHNSON (2012) CODES FOR REASONING-AND-PROVING TASKS

Type of RP	**Definition**	**Type of Case**	**Code**	**Connection to G. Stylianides' RP Framework (2010)**
Make a conjecture	Use a pattern to generate a conjecture	General case (nth term)	MG	Make a Conjecture
		Specific case (100th term)	MS	
Investigate a conjecture	Determine if a conjecture or assertion is true or false and provide a rationale	General case	IG	
		Specific case	IS	
Develop an argument	Write a proof of a statement (might have explain, explain why, show or show that)	General case	DG	Develop a proof (demonstration)
		Specific case	DS	Develop a proof (generic argument)
Evaluate an argument	Determine whether a stated argument is valid or not	General case	EG	Develop a proof (demonstration)
		Specific case	ES	Develop a proof (generic argument)

Counterexample	Find a counterexample/ disprove a statement		CX	Develop a proof
Correct or identify a mistake	A mistake is presented and student is asked to determine the error in reasoning.	General case	CG	Develop a non-proof argument
		Specific case	CS	
Principles of Proof	Explain how to outline an argument of a particular type, but not write a full proof.		PP	Develop an argument (proof or non-proof)

Appendix D

EXAMPLES OF THOMPSON, SENK, AND JOHNSON (2012) CODES FOR REASONING-AND-PROVING TASKS

Taken from Johnson, G. J., Thompson, D. R., & Senk, S. L. (2010). Proof-Related reasoning in high school textbooks. *Mathematics Teacher, 103* (6), 411-418.

Make a conjecture

Powers of $i, i^2, i^3, i^4, i^5, i^6, i^7$ ...have an interesting property. Using the fact that $i^2 = -1$, rewrite each power of i in simplest form and look for a pattern that would allow you to quickly rewrite powers like i^{100} or i^{523}.

(Core Plus, Course 4, 2001, p. 401)

Investigate a conjecture

You decide. Brittany tells Isabel that if $x + 3$ is a factor of the polynomial function $f(x)$, then $f(3) = 0$. Isabel argues that if $x + 3$ is a factor of $f(x)$, then $f(-3) = 0$. Who is correct? Explain.

(Glencoe, *Precalculus*, 2004, p. 226)

Develop an argument

Extended Response. Write a convincing argument to show why $3^0 = 1$ using the following pattern. $3^5 = 243$, $3^4 = 81$, $3^3 = 27$, $3^2 = 9$, …

(Glencoe, *Algebra 1*, 2004, p. 423)

Evaluate an argument

An algebra class has this problem on a quiz: "Find the value of $2x^2$ when $x = 3$." Two students reasoned differently. Student 1: Two times three is six. Six squared is thirty-six. Student 2: Three squared is nine. Two times nine is eighteen. Who was correct and why?

(Key Curriculum Press, *Discovering Algebra,* 2007, p. 353)

Counterexample

If c is a real number and n an odd positive integer, give an example to show that $x + c$ may not be a factor of $x^n - c^n$.

(Holt, Rinehart, and Winston, *Precalculus*, 2004, p. 250)

From *Precalculus*. Copyright 2004 by Holt, Rinehart, and Winston. All rights reserved. Reprinted by permission of the Houghton Mifflin Harcourt Publishing Company.

Correct or identify a mistake

Find the error in the following "proof" that $5 < 2$. Proof:

$$\frac{1}{32} < \frac{1}{4}$$

$$log\frac{1}{32} < log\frac{1}{4}$$

$$log\left[\left(\frac{1}{2}\right)^5\right] < log\left[\left(\frac{1}{2}\right)^2\right]$$

$$5log\left(\frac{1}{2}\right) < 2log\left(\frac{1}{2}\right)$$

$$5 < 2$$

(UCSMP, *Advanced Algebra*, 1996, p. 592)

Appendix E

TASK ANALYSIS GUIDE (STEIN ET AL., 2000)

Low-Level Cognitive Demands	High-Level Cognitive Demands
Memorization Tasks • Involve either producing previously learned facts, rules, formulae, or definitions *or* committing facts, rules, formulae, or definitions to memory. • Cannot be solved using procedures because a procedure does not exist or because the time frame in which the task is being completed is too short to use a procedure. • Are not ambiguous—such tasks involve exact reproduction of previously seen material and what is to be reproduced is clearly and directly stated. • Have no connection to the concepts or meaning that underlay the facts, rules, formulae, or definitions being learned or reproduced. *Procedures Without Connections Tasks* • Are algorithmic. Use of the procedure is either specifically called for or its use is evident based on prior instruction, experience, or placement of the task. • Require limited cognitive demand for successful completion. There is little ambiguity about what needs to be done and how to do it. • Have no connection to the concepts or	*Procedures With Connections Tasks* • Focus students' attention on the use of procedures for the purpose of developing deeper levels of understanding of mathematical concepts and ideas. • Suggest pathways to follow (explicitly or implicitly) that are broad general procedures that have close connections to underlying conceptual ideas as opposed to narrow algorithms that are opaque with respect to underlying concepts. • Usually are represented in multiple ways (e.g., visual diagrams, manipulatives, symbols, problem situations). Making connections among multiple representations helps to develop meaning. • Require some degree of cognitive effort. Although general procedures may be followed, they cannot be followed mindlessly. Students need to engage with the conceptual ideas that underlie the procedures in order to successfully complete the task and develop understanding. *Doing Mathematics Tasks* • Require complex and non-algorithmic thinking (i.e., there is not a predictable,

meaning that underlie the procedure being used. • Are focused on producing correct answers rather than developing mathematical understanding. • Require no explanations or explanations that focus solely on describing the procedure that was used.	well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example). • Require students to explore and to understand the nature of mathematical concepts, processes, or relationships. • Demand self-monitoring or self-regulation or one's own cognitive processes. • Require students to access relevant knowledge in working through the task. • Require students to analyze the task and actively examine task constraints that may limit possible solution strategies and solutions. • Require considerable cognitive effort and may involve some level of anxiety for the student due to the unpredictable nature of the solution process required.

Appendix F

DIRECTIONS FOR TASK COLLECTION

- Identify a unit in which students will have opportunities to develop their reasoning-and-proving understanding and skills between the beginning of February and the end of April. If the unit is longer than three weeks, identify a 15-day consecutive stretch of days to use for this study.
- Please submit all of the mathematical tasks you use for any purpose during the 15 consecutive days of instruction. Please also indicate which tasks are reasoning-and-proving mathematical tasks.
    - A photocopy of a textbook page with the tasks used is fine.
    - "Mathematical tasks" include any mathematical problems, exercises, examples, or individual or group work that students encounter, either in class or out of class.
    - If you modified any task, please provide both the original task and the modified task, and describe your rationale for modifying the task.
    - If you pulled a task from another source, identify the source.
- Please place the copies of the tasks in the file marked for the appropriate day, separating class work from homework. For each day, number the tasks according to their order in the day's lesson. On the log sheet provided in each day's folder, indicate the source of the task, approximately how much time was spent on the task and what purpose the task served in the lesson. For example, the task might have been used:
    - As a "warm-up" or "problem of the day"
    - To introduce the math ideas in the day's lesson
    - As independent or group work during class
    - As a homework assignment

Appendix G

REASONING-AND-PROVING TASK LOG SHEET

Day _______________ Teacher's Initials: _________________

"RP" means "reasoning-and-proving" and "HW" means "homework"

TASK #	SOURCE of the task	Was the task MODIFIED?	Is this a RP Task?	TIME SPENT on the task (or label HW)	PURPOSE of the task in the lesson

Appendix H

MODIFYING TASKS

Modification codes for tasks that the participant assigned an identified as reasoning-and-proving

Original Task (any source)	Task as assigned by teacher	Code
RP task	Task assigned, modified to LOWER RP	-1
RP task	Task assigned, NEUTRAL effect of modification	0
RP task	Task assigned, modified to INCREASE RP	+1
Non-RP task	Task assigned, modified to INCLUDE RP	+2

Reasoning-and-proving activities follow a hierarchy, from lowest to highest (G. Stylianides, 2010): Identify a pattern….make a conjecture…..provide a rationale….construct a proof (show "always true")

- A reasoning-and-proving textbook task could be modified by the teacher to <u>decrease</u> the reasoning-and-proving potential (e.g., written as make a conjecture, modified to only identify a pattern).

- Similarly, a reasoning-and-proving textbook task could be modified by the teacher to <u>increase</u> the reasoning-and-proving potential (e.g., written as make a conjecture, modified to construct a proof).

Appendix I

DIRECTIONS FOR STUDENT WORK (CLASSROOM ARTIFACT PACKET) COLLECTION

- Collect 3 class-sets of student work on reasoning-and-proving tasks. At least one of the tasks must require the students to show that something is "always true" (to write a proof). The "student work" is the written work from each student or group of students. Please do not include students' tests or quizzes.
- Please make copies of the students' work with the students' names removed (you can cut off the corners of the papers with their names).
- Please make a copy of the task as it was presented to the students.
- Complete a Student Work Cover Sheet for each class-set of student work.
- From each class-set of student work, identify:
    - 2 samples of work that exceeded expectations (mark with the GREEN stickers)
    - 2 samples of work that met expectations (mark with the YELLOW stickers)
    - 2 samples of work that failed expectations (mark with the RED stickers)
- Please place each set of student work, the task, and the Student Work Cover Sheet in the files marked for Student Work 1, Student Work 2, and Student Work 3.

Appendix J

STUDENT WORK (CLASSROOM ARTIFACT PACKET) COVER SHEET

Task # _______ on Day ________

1. Indicate if this assignment is **typical** (yes / no). If not, please explain:

2. Describe any directions—oral or written—you gave to the students that are not included on the task itself. Please explain any expectations you relayed to your class (e.g., work in groups, **expectations for grading**).

3. Did you implement the task differently than you had planned? If so, what changes did you make and why? What, if anything, surprised you during the enactment?

4. Explain your overall reaction to your implementation of this task (what do you believe your students and you learned, would you use this task again, etc.).

5. How did you assess students' work on the task? Please attach the specific or general criteria that you used.

Appendix K

IQA MATHEMATICS TOOLKIT RUBRICS

RUBRIC 1: Potential of the Task

4	**The task has the potential to engage students in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. **The task must explicitly prompt for evidence of students' reasoning and understanding.** For example, the task MAY require students to: Solve a genuine, challenging problem for which students' reasoning is evident in their work on the task; Develop an explanation for why formulas or procedures work; **Identify patterns and form generalizations based on these patterns;** **Make or investigate conjectures and support conclusions with mathematical evidence and/or create a proof or find a counterexample** **Evaluate an argument or explain how to outline an argument of a particular type.** Make explicit connections between representations, strategies, or mathematical concepts and procedures. Follow a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship

3	**The task has the potential to engage student in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the task does not warrant a "4" because:** The task does not explicitly prompt for evidence of students' reasoning and understanding. Students may be asked to engage in doing mathematics or procedures with connections, but the underlying mathematics in the task is not appropriate for the specific group of students (i.e., too easy or too hard to promote engagement with high-level cognitive demands); **Students may need to identify patterns but are not pressed for generalizations;** **Students may need to make conjectures but are not asked to support conclusions with mathematical evidence or create a proof** Students may be asked to use multiple strategies or representations but the task does not explicitly prompt students to develop connections between them; Students may be asked to make conjectures but are not asked to provide mathematical evidence or explanations to support conclusions
2	The potential of the task is limited to engaging students in using a procedure that is either specifically called for or its use is evident based on prior instruction, experience, or placement of the task. **There is little ambiguity about what needs to be done and how to do it.** The task does not require students to make connections to the concepts or meaning underlying the procedure being used. Focus of the task appears to be on producing correct answers rather than developing mathematical understanding (e.g., applying a specific problem solving strategy, practicing a computational algorithm). OR The task does not require students to engage in cognitively challenging work; the task is easy to solve.
1	**The potential of the task is limited to engaging students in memorizing or reproducing facts, rules, formulae, or definitions.** The task does not require students to make connections to the concepts or meaning that underlie the facts, rules, formulae, or definitions being memorized or reproduced. OR The task requires no mathematical activity.

RUBRIC 2: IMPLEMENTATION OF THE TASK

4	**Students engaged in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. **There is explicit evidence of students' reasoning and understanding.** For example, students may have: Solved a genuine, challenging problem for which students' reasoning is evident in their work on the task; Developed an explanation for why formulas or procedures work; **Identified patterns and formed generalizations based on these patterns;** **Made or investigated a conjectures and supported conclusions with mathematical evidence and/or created a proof or found a counterexample** **Evaluated an argument or explained how to outline an argument of a particular type.** Made explicit connections between representations, strategies, or mathematical concepts and procedures. Followed a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship
3	**Students engaged in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the implementation does not warrant a "4" because:** **There is no explicit evidence of students' reasoning and understanding.** Students engaged in doing mathematics or procedures with connections, but the underlying mathematics in the task was not appropriate for the specific group of students (i.e., too easy <u>or</u> too hard to sustain engagement with high-level cognitive demands); **Students identified patterns but did not make generalizations;** **Students made conjectures but did not provide sufficient mathematical evidence or explanations to support conclusions (e.g., students provided an non-proof argument in the form of an empirical argument or rationale)** Students used multiple strategies or representations but connections between different strategies/representations were not explicitly evident;
2	**Students engaged in using a procedure that was either specifically called for or its use was evident based on prior instruction, experience, or placement of the task. There was little ambiguity about what needed to be done and how to do it.** Students did not make connections to the concepts or meaning underlying the procedure being used. Focus of the implementation appears to be on producing correct answers rather than developing mathematical understanding (e.g.,, applying a specific problem solving strategy, practicing a computational algorithm).

	OR Students did not engage in cognitively challenging work; the task was easy to solve.
1	**Students engaged in memorizing or reproducing facts, rules, formulae, or definitions.** Students do not make connections to the concepts or meaning that underlay the facts, rules, formulae, or definitions being memorized or reproduced. OR Students did not engage in mathematical activity.

RUBRIC 5: CLARITY AND DETAIL OF EXPECTATIONS

4	The expectations for the quality of students' work are very clear and elaborated. Each dimension or criterion for the quality of students' work is clearly articulated. Additionally, varying degrees of success are clearly differentiated. **For a proof, the following criteria was used:** **The argument must show that the conjecture or claim is (or is not) true for *all* cases.** **The statements and definitions that are used in the argument must be ones that are true and accepted by the community because they have been previously justified.** **The conclusion that is reached from the set of statements must follow logically from the argument made.** The validity of the proof should NOT depend on: Type of proof (e.g., demonstration, generic example, exhaustion, induction) Form of the proof (e.g., two-column, paragraph, flow chart) Representation used (e.g., symbols, pictures, words) Explanatory power (e.g., how well the proof itself serves to explain why the claim is true)
3	**The expectations for the quality of students' work are clear and somewhat elaborated.** Levels of quality may be vaguely differentiated for each criterion (i.e., little information is provided for what distinguishes high, medium, and low performance).
2	**The expectations for the quality of students' work are broadly stated and unelaborated.**
1	**The teacher's expectations for the quality of students' work are unclear.** OR The expectations for quality work are not shared with students.

RUBRIC 6: COMMUNICATIONS OF EXPECTATIONS

4	Teacher **discusses the expectations** or criteria for student work (e.g., scoring guide, rubric) with students in advance of their completing the assignment **and models high-quality work.**

3	Teacher **discusses the expectations** or criteria for student work (e.g., scoring guide, rubric) with students in advance of their completing the assignment.
2	Teacher **provides a copy of the criteria** for assessing student work (e.g., scoring guide, rubric) to students in advance of their completing the assignment.
1	Teacher **does not share the criteria** for assessing students' work (e.g., scoring guide, rubric) with the students in advance of their completing the assignment (the teacher may provide a copy of the scoring rubric to students when giving them their final grade).
N/A	Reason:

Appendix L

SAMPLES OF TASKS FOR IQA RUBRIC 1: POTENTIAL OF THE TASK

Code 4: A Sticky Gum Problem

A Sticky Gum Problem

Ms. Hernandez came across a gumball machine one day when she was out with her twins. Of course, the twins each wanted a gumball. What's more, they insisted on being given gumballs of the same color. The gumballs were a penny each, and there would be no way to tell which color would come out next. Ms. Hernandez decides that she will keep putting pennies until she gets two gumballs that are the same color. She can see that there are only red and white gumballs in the machine.

1.) Why is three cents the most she will have to spend to satisfy her twins?
2.) The next day, Ms. Hernandez passes a gumball machine with red, white, and blue gumballs. How could Ms. Hernandez satisfy her twins with their need for the same color this time? That is, what is the most Ms. Hernandez might have to spend that day?
3.) Here comes Mr. Hodges with his triplets past the gumball machine in question 2. Of course, all three of his children want to have the same color gumball. What is the most he might have to spend?
4.) Generalize this problem as much as you can. Vary the number of colors. What about different size families? Prove your generalization to show that it always works for any number of children and any number of gumball colors.

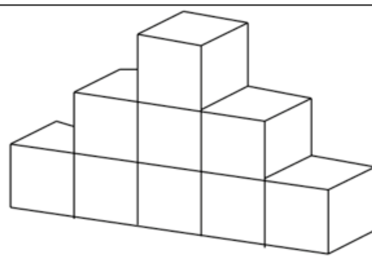
Fendel, D., Resek, D., Alper, L., & Frazer, S. (1996). *Interactive Mathematics Program Year 1—Unit 2: The Book of Pig* (p. 99). Emeryville, CA: Key Curriculum Press.

Comments: The students are provided with scaffolding that will help them identify a pattern, and they are asked to form a generalization based on that pattern. Additionally, the students are asked to construct a proof that shows their generalization holds for any number of children and any number of gumball colors.

Code 3: Toy Stack

TOYS Jamila is making a triangular wall with building blocks. The top row has one block, the second row has three, the third has five, and so on. How many rows can she make with a set of 100 blocks?

(Glencoe (2005), Algebra 2, pg. 586, #27)



Comments: A student may need to identify a pattern in order to determine the number of rows that can made with 100 blocks, but the student is not pressed for a generalization for any number of blocks nor is the student pressed to justify an answer with mathematical evidence.

Code 2: Simplifying an Expression

Simplify [the] expression, assuming that no variable equals zero. Write your answers with positive exponents only:

$$\left(\frac{1}{4}\right)^{-1}$$

(Holt, 2004, *Algebra 2*, pg. 99, #27)

From *Algebra 2*. Copyright 2004 by Holt, Rinehart, and Winston. All rights reserved. Reprinted by permission of the Houghton Mifflin Harcourt Publishing Company.

Comments: On page 94 of the text, the students were presented with the following "definition" of a negative exponent: "If n is a natural number, then $a^{-n} = \frac{1}{a^n}$". Thus, students only have to apply this procedure to complete this task. The students are not asked to explain why the negative exponent flips the fraction or provide any other evidence that their answer is correct.

Code 1: Name that Property

> State the property that is illustrated in the statement. All variables represent real numbers:
>
> $$4yw = 4wy$$
>
> (Holt, 2004, *Algebra 2*, p. 81, #54)

From *Algebra 2*. Copyright 2004 by Holt, Rinehart, and Winston. All rights reserved. Reprinted by permission of the Houghton Mifflin Harcourt Publishing Company.

Comments: Since the students were presented with a list properties and matching examples on p. 87 of the textbook, there is little ambiguity about what needs to be done. Students just need to choose the correct property from the list. The focus of the task appears to be to choose correctly;

the students are not asked to explain the commutative property of multiplication or explain, for instance, why there is no commutative property of subtraction.

Appendix M

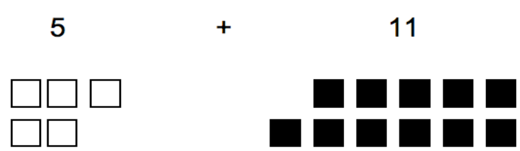
SAMPLES OF TASKS FOR IQA RUBRIC 1: IMPLEMENTATION OF THE TASK

The task used in this appendix was taken from the CORP course materials.

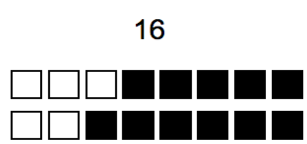
Task: *Prove that when you add any two odd numbers your answer is always even.*

Code 4

If I take the numbers 5 and 11 and organize the counters as shown, you can see the pattern.



You can see that when you put the sets together (add the numbers), the two extra blocks will form a pair and the answer is always even. This is because any odd number will have an extra block and the two extra blocks for any set of two odd numbers will always form a pair.

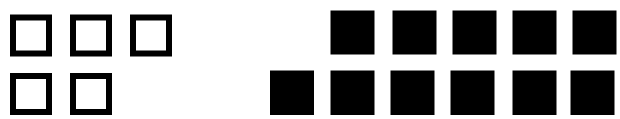


Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W. (2003). *Contemporary mathematics in context: A unified approach: Course 3.* New York, NY: Glencoe McGraw-Hill.

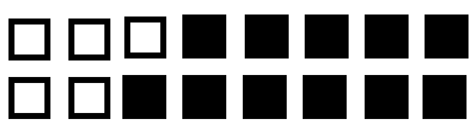
Comments: The students are clearly asked for a proof, and this student presented a general proof of the form generic example ("This is because **any** odd number will have an extra block…")

<u>Code 3</u>

If I take the numbers 5 and 11 and organize the counters as shown, you can see the pattern. 5 + 11



You can see that when you put the sets together (add the numbers), the two extra blocks will form a pair and the answer is always even.



Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W. (2003). *Contemporary mathematics in context: A unified approach: Course 3.* New York, NY: Glencoe McGraw-Hill.

Comments: Unlike the previous student, this student did not present a general argument; rather, the student presented an empirical rationale because the only numbers used were 5 and 11. The student did state that the answer will "always" be even, but it is unclear whether the sum of five and eleven will always be even or if the sum of two odds in general will always be even.

<u>Code 2</u>

My answer

Add 1 (a)	Add 2 (b)	a + b
1	3	4
7	9	16
11	13	24
21	23	44
113	97	210
1111	1111	2222
1003	10003	11006

I noticed all the sums will be an even number. a + b = c

Test: a = 35, b = 73 35 + 73 = 108 108 is also even so it is true.

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

Comment: The student was not held accountable for developing mathematical understanding; the student was allowed to produce a series of examples, present them in a way that does not clearly communicate thinking, and base the answer on one test case.

Code 1

$3a + 3b = 6(a + b)$

$a = 3$

$b = 9$

$(3 \times 3) + (3 \times 9) = 36$

$5a + 5b = 10(a + b)$

$93a + 57b = 140(a + b)$

An even number of odd numbers make an even answer but an odd number of odd numbers makes an odd answer:

Odd Even
$7a + 9b = 16(a + b)$

Odd Even Odd
$7a + 9b + 11c = 27(a + b + c)$

Odd Even Odd Even
$7a + 9b + 11c + 13d = 40(a + b + c + d)$

Odd Even Odd Even Odd

$93a + 7b + 13c + 101d + 39e = 153(a + b + c + d + e)$

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

Comments: It does not appear that the student engaged in mathematical activity. The student does not appear to have understood the task, nor are properties being used so the student's conclusion is invalid.

Appendix N

BACKGROUND INTERVIEW QUESTIONS

(Teachers were sent these questions prior to the first phone interview)

1. School descriptive information:
    a. Is your school urban, suburban, or rural?
    b. What grade levels are contained in your school?
    c. How many students are in your school?
    d. For what class did you select the unit for this study? How many students are in the class?
    e. How many classes and preparations do you have? How many total students do you have?
    f. How many math teachers work in your school?
2. Freedom to choose curriculum
    a. How do you select what you will teach (i.e. day-to-day curriculum or do you have flexibility in determining the curriculum for your students)?
    b. Is reasoning-and-proving an important feature of your curriculum? How often do you engage students in such activity?

c. How are you evaluated as a teacher? Who evaluates you?

3. Prior Reasoning-and-proving work with students

    a. Please describe any opportunities to develop students' understanding/skills for reasoning-and-proving you provided to your students prior to the unit you selected for this study.

    b. Do students have reference material (i.e. posters on the walls of the classroom, lists in their binders) available to them to help them engage in reasoning-and-proving tasks? If so, please send me pictures/copies of these materia

BIBLIOGRAPHY

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

Balacheff, N. (1987). Processes of proof and situations of validation. *Educational Studies in Mathematics, 18*(2), 147-176.

Balacheff, N. (1991). The benefits and limits of social interaction: The case of mathematical proof. In A. J. Bishop, E. Mellin-Olsen, & J. Van Dormolen (Eds.), *Mathematical knowledge: Its growth through teaching* (pp. 175-192). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Balacheff, N. (2010). Bridging knowing and proving in mathematics: a didactical perspective. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives.* New York: Springer.

Ball, D. L. (1988a). Research on teacher learning: Studying how teachers' knowledge changes. *Action in Teacher Education 10*(2), 17-23.

Ball, D. L. (1988b). Unlearning to teach mathematics. *For the Learning of Mathematics, 8*(1), 40-48.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is: Or might be: The role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6-8,14.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Towards a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco, Jossey-Bass.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59* (5), 389-407.

Barnett, C. (1991). Building a case-based curriculum to enhance the pedagogical content knowledge of mathematics teachers. *Journal of Teacher Education, 42*(4), 263-272.

Battista, M. T., & Clements, D. H. (2000). Mathematics curriculum development as a scientific endeavor. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 737-760). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Bell, A. W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics, 7*(1/2), 23-40.

Benson, J. et al (Eds.) (2009). *CME Project: Geometry* (Unit 2). Boston: Pearson.

Bieda, K. N. (2010). Enacting proof-related tasks in middle school mathematics: Challenges and opportunities. *Journal for Research in Mathematics Education, 41*(4), 351-382.

Blanton, & Stylianou (2002, October). Exploring sociocultural aspects of undergraduate students' transition to mathematical proof. In D. Mewborn, P. Sztajn, D. Y. White, H. G. Weigel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the Annual Meeting [of the] North American Chapter of the International Group for the Psychology of Mathematics Education: Vol*. 4 (p. 1671-1680). Retrieved from ERIC database (EJ471747).

Bogdan, R. C. & Biklen, S. K. (2003). *Qualitative research for education: An introduction to theory and methods.* Boston, MA: Allyn and Bacon.

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

Borko, H., Stecher, B. M., Alonzo, A. C., Moncure, S., & McClam, S. (2005). Artifact packages for characterizing classroom practice: A pilot study. *Educational Assessment, 10*(2), 73-104.

Boston, M. D. (2006). *Developing secondary mathematics teachers' knowledge of and capacity of implementing instructional tasks with high-level cognitive demands* (Doctoral dissertation, University of Pittsburgh). *Dissertation Abstracts International* (Publication No. AAT 3223943).

Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal, 113*(1), 76-104.

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119-156.

Boston, M., D., & Wolf, M. K. (2006). *Assessing academic rigor in mathematics instruction: The development of the Instructional Quality Assessment Toolkit.* (CSE Tech. Rep. No. 672). Los Angeles: University of California, Los Angeles, Center for the Study of Evaluation.

Boyd, C. J., Cummins, J., Malloy, C., Carter, J., Flores, A. (2005). *Geometry: Virginia Edition.* New York: Glencoe.

Boyer, C. B. (1991). *A history of mathematics* (2nd ed.). New York: John Wiley & Sons, Inc.

Brown, M., & Edelson, D. C. (2003, May). Teaching as design: Can we better understand the ways in which teachers use materials so we can better design materials to support their changes in practice? (LeTUS Report Series, RS-03). Evanston, IL: Northwestern University, The Center for Learning Technologies in Urban Schools.

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively-Guided Instruction: A knowledge base for reform in primary mathematics instruction. *The Elementary School Journal, 97*(1), 3-20.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: an experimental study. *American Educational Research Journal, 26*(4), 499-531.

Chapin, S. H., O'Connor, C. & Anderson, N. C. (2003). *Classroom discussions: Using math talk to help students learn, grades 1-6* (Chapter 1). Sausalito, CA: Math Solutions.

Charles, R. I., Hall, B., Kennedy, D., Bellman, A. E., Bragg, S. C., Handlin, W. G., Murphy, S. J., Wiggins, G. (2012). *Algebra 1: Common Core.* Boston: Pearson.

Charles, R. I., Hall, B., Kennedy, D., Bellman, A. E., Bragg, S. C., Handlin, W. G., Murphy, S. J., Wiggins, G. (2012). *Prentice Hall Algebra 1: Virginia Edition.* Boston: Pearson.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24(*4), 359-387.

Chazan, D., & Ball, D. L. (1999). Beyond being told not to tell. *For the Learning of Mathematics, 19*(2), 2-10.

Chazan, D., & Lueke, H. M. (2009). Exploring relationships between disciplinary knowledge and school mathematics: Implications for understanding the place of reasoning and proof in school mathematics. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and Learning Proof Across the Grades: A K-1 Perspective* (pp. 21-39). New York: Routledge.

Clare, L., & Aschbacher, P. R. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment, 7*(1), 39-59.

Clement, L. L. (2004). A model for understanding, using, and connecting representations. *Teaching Children Mathematics,11*(2), 97-102.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991).

Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*(1), 3-29.

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal, 20*(1), 41-53.

Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Hart, E. W., Keller, B. A., Watkins, A. E., Ritsema, B. E., & Walker, R. K. (2001). *Contemporary mathematics in context: A unified approach: Course 4.* Chicago: Everyday Learning.

Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W. (2003). *Contemporary mathematics in context: A unified approach: Course 3.* New York, NY: Glencoe McGraw-Hill.

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

Davis, P. J., & Hersh, R. (1981). *The mathematical experience.* Boston: Houghton Mifflin.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3-14.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53*(2), 159-199.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*(2), 167-180.

de Villiers, M. (1999). *Rethinking proof with the Geometer's Sketchpad.* Emeryville, CA: Key Curriculum Press.

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.

Edwards, L. D. (1999). Odds and evens: Mathematical reasoning and informal proof among high school students. *Journal of Mathematical Behavior, 17*(4), 489-504.

Ellis, A. B. (2011). Generalizing-promoting actions: How classroom collaborations can support students' mathematical generalizations. *Journal for Research in Mathematics Education, 42*(4), 308-345.

English, L., D., & Warren, E. A. (1995). General reasoning processes and elementary algebraic understanding: Implications for initial instruction. *Focus on Learning Problems in Mathematics, 17*(4), 1-19. Retrieved from http:eric.ed.gov (EJ526526)

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119-161). New York, NY: MacMillan Press.

Fendel, D., Resek, D., Alper, L., & Frazer, S. (1996). *Interactive Mathematics Program Year 1—Unit 2: The Book of Pig* (p. 99). Emeryville, CA: Key Curriculum Press

Fennema E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use childrens' thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

Fennema, E., Franke, M. L., Carpenter, T. P., & Carey, D. A. (1993). Using children's mathematical knowledge in instruction. *American Educational Research Journal, 30*(3), 555-583.

Fischbein, E., (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9-18, 24.

Fischbein, E. (1999). Intuitions and schemata in mathematical reasoning. *Educational Studies in Mathematics, 38*(1/3), 11-50.

Friel, S. N., & Carboni, L. W. (2000). Using video-based pedagogy in an elementary mathematics methods course. *School Science and Mathematics, 100*(3), 118-27.

Fuson, K. C., & Briars, D. J. (1990). Using a base-ten blocks learning/teaching approach for first- and second-grade place-value and multidigit addition and subtraction. *Journal for Research in Mathematics Education, 21*(3), 180-206.

Galbraith, P. (1981). Aspects of proving: A clinical investigation of process. *Educational Studies in Mathematics, 12* (1), 1-28.

Goetting, M. M. (1995). *The college students' understanding of mathematical proof* (Doctoral dissertation, University of Maryland College Park). Available from ProQuest Dissertations and Theses database. (UMI No. 304216478)

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

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009a). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2055-2100.

Grossman, P., Hammerness, K., & McDonald, M. (2009b). Redefining teaching, re-imagining teacher education. *Teachers and Teaching: theory and practice, 15*(2), 273-289.

Grouws, D. A., & Smith, M. S. (2000). NAEP findings on the preparation and practices of mathematics teachers. In E. A. Silver & Kenney, P. A. (Eds.), *Results from the seventh mathematics assessment of the National Assessment of Educational Progress* (pp. 107-140). Reston, VA: National Council of Teachers of Mathematics.

Grouws, D. A., Smith, M. S., & Sztajn, P. (2004). The preparation and teaching practices of United States mathematics teachers: Grades 4 and 8. In Kloosterman, P. & Lester, F. K. (Eds.) *Results and Interpretations of the 1990 through 2000 Mathematics Assessments of the National Assessment of Educational Progress,* Reston, VA: NCTM.

Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics, 44*(1/2), 127-150.

Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal, 28*(4), 567-590.

Hanna, G. (1983). *Rigorous proof in mathematics education.* Ontario Institute for Studies in Education Curriculum Series: Vol. 48. Toronto, Ontario: OISE Press.

Hanna, G., & Jahnke, H. N. (1993). Proof and application. *Educational Studies in Mathematics, 24*(4), 421-438.

Harel, G., & Rabin, J. M. (2010). Teaching practices associated with the authoritative proof scheme. *Journal for Research in Mathematics Education, 41*(1), 14-19.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *CMBS Issues in Mathematics Education: Vol. 7 Research in Collegiate Mathematics Education III (pp.* 234-283). Washington, DC: American Mathematical Society.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 805-842). Charlotte, NC: Information Age.

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

Healy, L., & Hoyles, C. (2001). Software tools for geometrical problem solving: Potentials and pitfalls. *International Journal of Computers for Mathematical Learning, 6*, 235-256.

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

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

Hiebert, J., Carpenter, T. P., Fenema, E., Fuson, K. C., Wearne, D., Murray, N., et al. (1997).

Making sense: Teaching and learning mathematics with understanding. Portsmouth, NH: Heinemann.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*(2), 393-425.

Holliday, B., Cuevas, G. J., McClure, M. S., Carter, J. A., & Marks, D. (2004). *Advanced mathematical concepts: Precalculus with applications.* Columnbus, OH: Glencoe/McGraw-Hill.

Holliday, B., Cuevas, G. J., Moore-Harris, B., Carter, J. A., Marks, D., Casey, R. M., Day, R., & Hayek, L. M. (2004). *Algebra 1.* Columbus, OH: Glencoe/McGraw-Hill.

Horizon Research, Inc. (2003). *Looking inside the classroom: A study of K-12 mathematics and science education in the United States.* Chapel Hill, NC: Horizon Research, Inc.

Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics, 17*(1), 7-16.

Hungerford, T. W., Jovell, I., & Mayberry, B. (2004). *Precalculus: A Graphing Approach.* Austin, TX: Holt, Rinehart, and Winston.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*(4), p. 358-390.

Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: a typology of reflective practice for teacher education. *Teaching and Teacher Education, 18*, 73-85.

Johnson, G. J., Thompson, D. R., & Senk, S. L. (2010). Proof-related reasoning in high school textbooks. *Mathematics Teacher, 103*(6), 441-418.

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

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

Knuth, E. J. (2002c). Teachers' conceptions of proof in the context of secondary school Mathematics. *Mathematics Teacher Educator, 5,* 61-88.

Knuth, E. J., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Weigel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24th Annual Meeting of the North American Chapter of the International Group for the*

Psychology of Mathematics Education (Vol. 4, pp. 1693-1670). Athens, GA: Clearinghouse for Science, Mathematics, and Environmental Education.

Knuth, E. J., & Sutherland, J. (2004). Student understanding of generality. In D. McDougall (Ed.), *Proceedings of the 26th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp.561-567). Toronto: University of Toronto/OISE.

Ko, Y., & Knuth, E. (2009). Undergraduate mathematics majors' writing performance producing proofs and counterexamples about continuous functions. *Journal of Mathematical Behavior, 28*, 68-77.

Konior, J. (1993). Research into the construction of mathematical texts. *Educational Studies in Mathematics, 24*(3), 251-256.

Küchemann, D., & Hoyles, C. (2001, July). Investigating factors that influence students' mathematical reasoning. In M. van den Heuvel-Panhuizen (Ed.) *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education: PME 25@NL; 12-17 July, 2001.* Ultrecht, The Netherlands: Freudenthal Institute, Faculty of Mathematics and Computer Science, Utrecht University.

Laborde, C. (2000). Dynamic geometry environments as a source of rich learning contexts for the complex activity of proving. *Educational Studies in Mathematics, 44* (1/2), 151-161.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery.* Cambridge: Cambridge University Press.

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

Lappan, G., & Briars, D. (1995). "How Should Mathematics be Taught?" In I. Carl (Ed.), *Prospects for school mathematics*. Reston, Va.: National Council of Teachers of Mathematics.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Philips, E. D. (1998/2004). *Connected Mathematics Project.* Menlo Park, CA: Dale Seymour Publications.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Philips, E. D. (1996). *Comparing and scaling: Connected Mathematics Project.* Palo Alto, CA: Dale Seymour.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry.* Thousand Oaks, CA: Sage.

Lithner, J. (2004). Mathematical reasoning in calculus textbooks exercises. *Journal of Mathematical Behavior, 23,* 405-427.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 15*(2), 129-151.

Loucks-Horsely, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics, 99*(5), 258-271.

Love, E., & Pimm, D. (1996). 'This is so': a text on texts. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 371- 409). Boston: Kluwer Academic Publishers.

Manin, I., & Zilber, B. (2010). *A course in mathematical logic for mathematicians (2nd ed.).* New York: Springer.

Mariotti, M. A. (2000). Introduction to proof: The mediation of a dynamic software environment. *Educational Studies in Mathematics, 44*(1/2), 25-53.

Mariotti, M. A. (2001). Justifying and proving in the Cabri environment. *International Journal of Computers for Mathematical Learning, 6*, 257-281.

Marrades, R., & Guttierez, A. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. Educational Studies in Mathematics, 44, 87–125.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41-51.

Martin, T. S., McCrone, S. M. S., Bower, M. L. W., & Dindyal, J. (2005). The interplay of teacher and student actions in the teaching and learning of geometric proof. *Educational Studies in Mathematics, 60,* 95-124.

Matsumura, L. C., Slater, S. C., Junker, B., Peterson, M., Boston, M., Steele, M., Resnick, L. (2006). *Measuring reading comprehension and mathematics instruction in urban middle schools: A pilot study of the instructional quality assessment*. (CSE Tech. Rep. No. 681). Los Angeles: University of California, Los Angeles, Center for the Study of Evaluation.

Matsumura, L. C., Garnier, H., Pascal, J., & Valdés, R. (2002). Measuring instructional quality in accountability systems: Classroom assignments and student achievement. *Educational Assessment*, *8*(3)**,** 207–229.

Matsumura, L. C., Garnier, H., Slater, S. C., & Boston, M. D. (2008). Toward measuring instructional interactions "at-scale." *Educational Assessment, 13,* 267-300.

Mayer, R. E., Tajika, H., & Stanley, C. (1991). Mathematical problem solving in Japan and the United States: A controlled comparison. *Journal of Educational Psychology, 83*(1), 69-72.

McCrone, S. M. S., & Martin, T. S. (2009). Formal proof in high school geometry: Student perceptions of structure, validity, and purpose. In Stylianou, D. A., Blanton, M. L., and Knuth, E. J. (Eds.) *Teaching and Learning Proof Across the Grades: A K-16 Perspective* (pp. 204-221). New York: Routledge.

Merriam, S. B. (1998). *Qualitative research and case study application in education.* San Francisco, CA: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Thousand Oaks, CA: Sage.

Mingus, T. T., & Grassl, R. M. (1999). Preservice teacher beliefs about proofs. *School Science and Mathematics, 99*(8), 438-444.

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

Morris, A. K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction, 20*(1), 79-118.

Murdock, J., Kamischke, E., & Kamischke, E. (2007). *Discovering algebra: An investigative approach, 2nd ed.* Emeryville, CA: Key Curriculum Press.

National Commission on Excellence in Education. (1983. *A Nation at Risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

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

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

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics*. Washington, DC: Author.

Pegg, J., & Redden, E. (1990) Procedures for, and experiences in, introducing algebra in New South Wales. *Mathematics Teacher, 83*(5), 386-391.

Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics, 21,* 589-598.

Porter, A., (1993). *Opportunity to learn* (Brief No. 7). Madison, WI: Center on Organization and Restructuring of Schools.

Putnam, R. T. (1992). Teaching the "hows" of mathematics for everyday life: A case study of a fifth-grade teacher. *Elementary School Journal, 93*(2), 163-177.

Quint, J. C., Akey, T. M., Rappaport, S., & Willner, C. J. (2007). *Instructional leadership, teaching quality, and student achievement: Suggestive evidence from three urban school districts.* New York: MDRC. Retrieved from http://www.mdrc.org/sites/default/files/Instructional%20Leadership%20ES.pdf

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics, 52*(3), 319-325.

Raudenbush, S. W., Rowan, B., Cheong, Y. F. (1990). *Contextual effects on the self-efficacy of high school teachers*. Stanford, CA: Center for Research on the Context of Secondary Teaching, School of Education.

Remillard, J. T. (2000). Can curriculum materials support teachers' learning? Two fourth-grade teachers' use of a new mathematics text. *Elementary School Journal, 100*(4), 331-350.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211-246.

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

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

Sadri, P. (2008). *Teaching mathematics for understanding: Intentions and practices of an expert middle school mathematics teacher in the context of a reformed-based curriculum.* (Oregon State University). Available from *ProQuest Dissertations and Theses* (UMI No. 230667914).

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*(3), 299-311.

Sather, S. E. (1994). *Complexities of control: Teachers' influence over their own work in secondary schools.* (Doctoral dissertation, University of California, Berkeley). Available from *ProQuest Dissertations and Theses* (UMI No. 304083524)

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of U. S. science and mathematics education*. Hingham, MA: Kluwer.

Schneider, R., & Krajcik, J. (2002). Supporting science teacher learning: The role of educative curriculum materials. *Journal of Science Teacher Education, 13*(3), 221-245.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving.* Orlando, Florida: Academic Press.

Schoenfeld, A. H. (1986). On having and using geometric knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 225-264). Hillsdale, NJ: Erlbaum.

Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education, 20*(4), 338-355.

Schoenfeld, A. H., (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior, 13,* 55-80.

Schofield, J. W. (2002). Increasing the generalizability of qualitative research. In A. M. Huberman & M. B. Miles (Eds.), *The qualitative researcher's companion* (pp. 171-204). Thousand Oaks, CA: Sage.

Schultz, J. E., Ellis, W., Hollowell, K. A., Kennedy, P. A. (2004). *Algebra 2.* New York: Holt, Rinehart and Winston.

Seeley, C., (2003). Mathematics textbook adoption in the United States. In G. M. A. Stanic & J. Kilpatrick (Eds.), *A history of school mathematics* (pp. 957-989). Reston, VA: National Council of Teachers of Mathematics.

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. (2003, August). Errors and misconceptions in college level theorem proving (Report No. 2003-3). Cookeville, TN: Tennessee Technological University.

Senk, S. L. (1985). How well do students write geometry proofs? *Mathematics Teacher, 78*(6), 448-456.

Senk, S. L., Thompson, D. R., Viktora, S. S., Usiskin, Z., Ahbel, N. P., Levin, S., Weinhold, M. L., Rubenstein, R. N., Jaskowiak, J. H., Flanders, N. J., & Pillsbury, G. (1996). *Advanced Algebra, 2nd ed.* Glenview, I: Scott Foresman.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Silver, E. A., Clark, L. M., Ghousseini, H. N., Charalambous, C. Y., & Sealy, J. T. (2007). Where is the mathematics? Examining teachers' mathematical learning opportunities in practice-based professional learning tasks. *Journal of Mathematics Teacher Educators, 10*, 261-277.

Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Strawhun, B. T. F. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *Journal of Mathematical Behavior, 24,* 287-301.

Silver, E. A., Mesa, V., Morris, K., Star, J., & Benken, B. (2009). Teaching for understanding: An analysis of mathematics lessons submitted by teachers seeking NBPTS certification. *American Educational Research Journal, 46* (2), 501-531.

Silver, E. A., Mills, V., Castro, A., & Ghousseini, H. (2006) Blending elements of lesson study With case analysis and discussion: A promising professional development synergy. *AMTE Monograph 3: The Work of Mathematics Teacher Educators,* 117-132.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. Journal for Research in Mathematics Education, 26, 114-145.

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

Sleep, L. (2009). *Teaching to the mathematical point: Knowing and using mathematics in teaching.* (Doctoral dissertation, University of Michigan). Available from ProQuest Dissertations and Theses (UMI No. 304929452).

Smith, M. S. (2001). *Practice-based professional development for teachers of mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Smith, M. S., Arbaugh, F., Steele, M., Boyle, J., Fulkerson, W., Knouck, N., & Vrabel, A. (April, 2012). *Developing Teachers' Capacity to Support Students' Reasoning and Proving.* Presented at the annual meeting of the National Council of Teachers of Mathematics, Philadelphia, Pennsylvania.

Stacey, K., & MacGregor, M. (1997). Building foundations for algebra. *Mathematics Teaching in the Middle School, 2*(4)*,* 252-260.

Stein, M. K., Baxter, J. A., & Leinhardt, G. (1990). Subject-matter knowledge and elementary instruction: A case from functions and graphing. *American Educational Research Journal, 27*(4), 639-663.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313-340.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455-488.

Stein, M. K., & Kim, G., (2006, April). The role of mathematics curriculum in large-scale urban reform: An analysis of demands and opportunities. *Paper presented at the Annual Meeting of the American Educational Research Association,* San Francisco.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: an analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50-80. Retrieved from http://dx.doi.org/10.1080/1380361960020103.

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3,* 268-275.

Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2000, 2009). *Implementing standards-based mathematics instruction: A casebook for professional development.* New York, NY: Teacher's College Press.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: The Free Press.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*(3), 289-321.

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics, 55*, 133-162.

Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics, 72,* 237-253.

Stylianides, G. J. (2007). Investigating the guidance offered to teachers in curriculum materials: The case of proof in mathematics. *International Journal of Science and Mathematics Education, 6*(1), 191-215.

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 28 (*1), 9-16.

Stylianides, G. J. (2009). Reasoning-and-Proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11,* 258-288.

Stylianides, G. J. (2010). Engaging secondary students in reasoning-and-proving. *Mathematics Teaching,* September, 39-44.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*(3), 314-342.

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Educators, 10,* 145-166.

Stylianou, D., & Smith, M. S. (2000). Examining student responses: A strategy for developing preservice elementary teachers' understanding of algebra. In M. K. Heid, M. S. Smith, and G. W. Blume (Eds.) *Algebra across the grades, 2000 yearbook of the Pennsylvania Council of Teachers of Mathematics* (pp. 23-32). Pennsylvania Council of Teachers of Mathematics.

Szombathely, A., & Szarvas, T. (1998). Ideas for developing students' reasoning: A Hungarian perspective. *Mathematics Teacher, 91*(8), 677-681.

Tabach, M., Levenson, E., Barkai, R., Tsamir, P., Tirosh, D., & Dreyfus, T. (2011). Secondary teachers' knowledge of elementary number theory proofs: the case of general-cover proofs. *Journal of Mathematics Teacher Educators, 14*, 465-481.

Tarr, J. E., Chavez, O., Reys, R. E., & Reys, B. J. (2006). From the written to the enacted curricula: The intermediary role of middle school mathematics teachers in shaping students' opportunity to learn. *School Science and Mathematics, 106*, 191-201.

Thompson, A. G., & Thompson, P. W. (1996). Talking about rates conceptually, part 2: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 27(1),* 2-24.

Thompson, D. R., Senk, S. L., & Johnson, G. J. (2012). Opportunities to learn reasoning and Proof in high school mathematics textbooks. *Journal for Research in Mathematics Education, 43*(3), 253-295.

Thompson, C. L., & Zeuli, J. S. (1999). The frame and the tapestry: Standards-based reform and professional development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 341-375). San Francisco: Jossey-Bass.

Thurston, W. P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161-177.

Tinto, P. P. (1990). *Students' views on learning proof in high school geometry: An analytic-inductive approach* (Doctoral Dissertation, Syracuse University). Available from ProQuest Dissertations and Theses (UMI No. 303905361).

Tyson-Bernstein, H., & Woodward, A., (1991). Nineteenth century policies for twenty-first century practice: The textbook reform dilemma. In P. G. Altbach, G. P. Kelly, H. G. Petrie, & L. Weis (Eds.), *Textbooks in American society* (pp. 91-104). New York: State University of New York Press.

Van Dormolen, J. (1977). Learning to understand what giving a proof really means. *Educational Studies in Mathematics, 8*(1), 27-34.

Varghese, T. (2009). Secondary-level student teachers' conceptions of mathematical proof. *Issues in the Undergraduate Mathematics Preparation of School Teachers: The Journal, 1*. Retrieved from http://www.k-12prep.math.ttu.edu.

Wearne, D., & Hiebert, J. (1989). Cognitive changes during conceptually based instruction on decimal fractions. *Journal of Educational Psychology, 81*(4), 507-513.

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

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*(4), 431-459.

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

Williams, E. R. (1979). *An investigation of senior high school students understanding the nature of mathematical proof* (Doctoral dissertation, University of Alberta Edmonton). Available from ProQuest Dissertations and These database. (UMI No. 302978354)

Wu, H.-H. (1996). The role of Euclidean geometry in high school. *Journal of Mathematical Behavior, 15*(3), 221-237.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458-477.

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp. 227-236). Reston, VA: The National Council of Teacher of Mathematics.

Yin, R. K. (1993). *Applications of case study research.* Thousand Oaks, CA: Sage.

Zeichner, K. M., & Liston, D. P. (1996). *Reflective teaching: An introduction.* Mahwah, New Jersey: Lawrence Erlbaum Associates.